

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6: C12N 15/11, 15/00, 15/63, C07H 21/02, 21/04		A1	(11) International Publication Number: WO 00/04140 (43) International Publication Date: 27 January 2000 (27.01.00)									
<p>(21) International Application Number: PCT/US99/15849</p> <p>(22) International Filing Date: 14 July 1999 (14.07.99)</p> <p>(30) Priority Data:</p> <table> <tr> <td>60/092,921</td> <td>15 July 1998 (15.07.98)</td> <td>US</td> </tr> <tr> <td>60/092,922</td> <td>15 July 1998 (15.07.98)</td> <td>US</td> </tr> <tr> <td>60/092,956</td> <td>15 July 1998 (15.07.98)</td> <td>US</td> </tr> </table> <p>(71) Applicant (<i>for all designated States except US</i>): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (<i>for US only</i>): RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). KOMATSOULIS, George [US/US]; 9518 Garwood Street, Silver Spring, MD 20901 (US). DUAN, Roxanne, D. [US/US]; 5515 Northfield Road, Bethesda, MD 20817 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). MOORE, Paul, A. [US/US]; 19005 Leatherbark Drive, Germantown, MD 20874 (US). SHI, Yang-gu [CN/US]; Apartment 102, 437 West Side Drive, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 3142 Quesada Street, N.W., Washington, DC 20015 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne</p>				60/092,921	15 July 1998 (15.07.98)	US	60/092,922	15 July 1998 (15.07.98)	US	60/092,956	15 July 1998 (15.07.98)	US
60/092,921	15 July 1998 (15.07.98)	US										
60/092,922	15 July 1998 (15.07.98)	US										
60/092,956	15 July 1998 (15.07.98)	US										
<p>Terrace #316, Gaithersburg, MD 20878 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). BREWER, Laurie, A. [US/US]; Apartment 115, 410 Van Dyke Street, St. Paul, MN 55119-4321 (US). FLORENCE, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). MUCENSKI, Michael [US/US]; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). SOPPET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, VA 22020 (US).</p> <p>(74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).</p> <p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published <i>With international search report.</i></p>												
<p>(54) Title: 71 HUMAN SECRETED PROTEINS</p> <p>(57) Abstract</p> <p>The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.</p>												

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

71 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and 5 polypeptides, and their production.

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or 10 organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic 15 reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

20 Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in 25 secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes 30 encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon, Factor VIII, human

growth hormone, tissue plasminogen activator, and erythropoietin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent

5 medical disorders by using secreted proteins or the genes that encode them.

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, 10 antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

15

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

20 In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of 25 matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released 30 into the extracellular space, the secreted protein can undergo extracellular processing

to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

- In specific embodiments, the polynucleotides of the invention are less than 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, or 7.5 kb in length. In a further 5 embodiment, polynucleotides of the invention comprise at least 15 contiguous nucleotides of the coding sequence, but do not comprise all or a portion of any intron. In another embodiment, the nucleic acid comprising the coding sequence does not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene in the genome).
- 10 As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the 15 secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X 20 was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, 25 Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those 30 polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an

overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

- 5 Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or
10 temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH₂PO₄; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent
15 hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTO, heparin, denatured salmon sperm DNA, and
20 commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a
25 complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The polynucleotide of the present invention can be composed of any
30 polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of

single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single-

- 5 and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of
- 10 modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids.

- 15 The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol,
- 20 25 30 cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation,

gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

- 5 (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

- 10 "SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

- 15 "A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not 20 more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

Polynucleotides and Polypeptides of the Invention

- 25 **FEATURES OF PROTEIN ENCODED BY GENE NO: 1**

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

- 30 PFCSGGFPSSLWIYLPFIFNVSDLWMGSLSGCALPFCLXVFFLTVPSPAVGLLXF
AGGPLQTLFAWVSPVEAAEQQRLLPVLSGSFVSEGTCQMPARALLYEVSVG

PYWEIPPSQDTRRSGYLRRQSDP (SEQ ID NO: 195) . Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in pancreas islet cell tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as

- 5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the pancreas, including cancer and diabetes. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For
- 10 a number of disorders of the above tissues or cells, particularly of the pancreas, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., endocrine, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the
- 15 standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- The tissue distribution in tumors of pancreatic islet cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis, treatment and intervention of such tumors, in addition to other endocrine or
- 20 gastrointestinal tumors where expression has been indicated. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or
 - 25 immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
- 30

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1099 of SEQ ID NO:11, b is an integer of 15 to 1113, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the

10 following amino acid sequence:

HEGSCRAPGFSAHKGRCPSRMTLPSRALASLGVGVWGMRLRNQVTVSCG
GSRWSSRVALGAFSWVCVALVLQPSGGGLGLTSPSEGCWEHELALAVLRA
PGGS (SEQ ID NO: 196) . Polynucleotides encoding these polypeptides are also provided.

15 This gene is expressed equally in in .

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic disorders, particularly leukemia. Similarly,

20 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular and immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hemolymphoid, 25 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 104 as residues: Gly-29 to Ser-35, Ser-63 to Cys-68. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in hemangiopericytoma, breast lymph node, and bone marrow indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 969 of SEQ ID NO:12, b is an

integer of 15 to 983, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

5 The translation product of this gene shares sequence homology with the *Drosophila melanogaster* slit protein, a secreted protein that contains both an EGF domain and Leucine Rich Repeat domains. It is thought to be important in the development of midline glia and commissural axon pathways (See e.g., Rothberg et al. *Genes Dev.* 4:2169-87 (1990); which is hereby incorporated by reference herein).

10 This gene is expressed primarily in human hippocampus. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, and developmental disorders. Similarly, polypeptides and 15 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neurological system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neurological, cancerous, or wounded tissues) or 20 bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution within human hippocampus combined with the homology to the *Drosophila* slit protein, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and 30 elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease,

Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning

- 5 disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. In addition, elevated expression of this gene product in regions
10 of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to

- 15 identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
20 related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
25 formula of a-b, where a is any integer between 1 to 959 of SEQ ID NO:13, b is an integer of 15 to 973, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

- 30 In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by

the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

IPLTLPGIFLLIRLFWRLGQSICPGKLVLWPQFCGGCAVISGHCVPRGMPSW
LPGCFVLLCLVAVGQLREWGVGGSAGVLLALPHLQVLGMRGRGLISGG

5 (SEQ ID NO: 197). Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 16. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 16.

10 This gene is expressed in KMH2 cells, osteoblasts, fetal spleen, Jurkat membrane bound polysomes, breast, and cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer, immune, and skeletal disorders. Similarly, polypeptides and 15 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, skeletal, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the 20 standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in KMH2 cells, osteoblasts, and fetal spleen indicates 25 that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product in fetal spleen and T-cells indicates a role in the regulation of the 30 proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved

in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

- Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, scleroderma and tissues. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1444 of SEQ ID NO:14, b is an

integer of 15 to 1458, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

- 5 The translation product of this gene shares sequence homology with phospholipase A2 which cleaves fatty acids from carbon 2 of glycerol (ref. Prosite pattern documentation for PS2_HIS). Many snake venoms contain phospholipase A2, which prevents transmission of nerve impulses to muscles by blocking the release of acetylcholine from the neuron. Therefore, included in this invention as preferred
- 10 domains are Phospholipase A2 histidine active site domains, which were identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). Phospholipase A2 is an enzyme which releases fatty acids from the second carbon group of glycerol. Structurally, PA2's are small and rigid proteins of 120 amino-acid residues that have four to seven disulfide bonds. PA2 binds a calcium ion which is required for activity.
- 15 The side chains of two conserved residues, a histidine and an aspartic acid, participate in a 'catalytic network'. Two different signature patterns for PA2's were developed. The first is centered on the active site histidine and contains three cysteines involved in disulfide bonds. The consensus pattern is as follows: C-C-x(2)-H-x(2)-C [H is the active site residue].
- 20 Preferred polypeptides of the invention comprise a Phospholipase A2 histidine active site domain selected from the following amino acid sequences: CCNQHDRC (SEQ ID NO: 199), SLTKCCCNQHDRCYET (SEQ ID NO: 200) , and/or LTKCCCNQHDRCYETCG (SEQ ID NO: 201) . Polynucleotides encoding these polypeptides are also provided. Further preferred are polypeptides comprising the
- 25 Phospholipase A2 histidine active site domain of the sequence listed in Table 1 for this gene, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of this referenced sequence. The additional contiguous amino acid residues is N-terminal or C-terminal to the Phospholipase A2 histidine active site domain. Alternatively, the additional contiguous amino acid residues is both N-terminal and
- 30 C-terminal to the Phospholipase A2 histidine active site domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The

above preferred polypeptide domain is characteristic of a signature specific to Phospholipase A2 proteins. Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with Phospholipase A2 proteins. Such activities are known in the art, some of which are described

- 5 elsewhere herein, or see, for example, McIntosh, et al. J. Biol. Chem. 270 (8), 3518-3526 (1995), incorporated herein by reference.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the

- 10 following amino acid sequence:

GPAGKEAWIWSWLLPSPGPAPLPSASWGLCGDAPR
AAARGPVEPGAARMALLSRPALTLRLMAAVVRCQEQAQTTDWWRATLKTI
RNGVHKIDTYLNAALDLGGEDGLCQYKCSDGSKPFPRYGYKPSPPNGCGSP
LFGXHLNIGIPSLTKCCNQHDRCYETCGKSKNDCDEEFQYCLSKICRDVQKTL
15 GLTQHVQACETTVELLFDSEHILGCKPYLDSQRAACRCHYEKSDL (SEQ ID NO: 198) . Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome 4. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 4.

- 20 This gene is expressed in a variety of cell types with no single type predominating.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are 25 not limited to, neurological disorders, or metabolism disorders, specifically phospholipase A2 deficiencies. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neuromuscular system, expression of this gene at 30 significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., pancreas, cancerous and wounded tissues) or bodily fluids (e.g., bile, lymph,

serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 107 as residues: Gln-23 to Asp-30, Lys-66 to Cys-87. Polynucleotides encoding said polypeptides are also provided.
- The ubiquitous tissue distribution and homology to phospholipase A2 indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis 10 and treatment of neuromuscular disorders. Alternatively, considering the activity of phospholipase A2 as a block for neuro- transmission may suggest that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the 15 gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Alternatively, the homology to 20 Phospholipase A2 proteins may indicate a potential use for the protein product of this gene in diagnosis, treatment and/or prevention of metabolism disorders, specifically deficiencies in Phospholipase A2. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate 25 ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly 30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

- 5 formula of a-b, where a is any integer between 1 to 1991 of SEQ ID NO:15, b is an integer of 15 to 2005, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

- 10 In a specific embodiment polypeptides of the invention comprise the following amino acid sequence:

GTSSARPRGALPGGSAPSAPHQLPGRAQAPAVSGPPPTSGLCHFDPAAPWPL
WPGPWQLPPHPQDWPAHPDIPQDWVSLRSFGQLTLCPRNGTVTGKWRGSH
VVGLLTTLNFGDGPDRNKTRTFQATVLGSQMGLKGSSAGQLVLITARVTTER

- 15 TAGTCLYFSAVPGILPSSQPPISCSEEGAGNATLSPRMGEECVSVWSHEGLVLT
KLLTSEELALCGSRLLVLGSFLLFCGLLCCVTAMCFHPRRESHWSRTRL
(SEQ ID NO: 202) . Polynucleotides encoding these polypeptides are also provided.

- In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by
20 the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

ARAPPGPEGLSPEAQPPLLPMGNCQAGHNLHLCLAHPPPLVCATLILLLGLS
GLGLGSFLLTHRTGLRT LTSPRTGSLF (SEQ ID NO: 203) . Polynucleotides
encoding these polypeptides are also provided.

- 25 This gene is expressed in a wide variety of tissue types including testes, cerebellum, dendritic cells, breast cancer, umbilical vein endothelial cells, epididymus, corpus colosum, chronic synovitis, liver hepatome, normal breast, osteoblasts, melanocytes, B cell lymphomas, and to a lesser extent in other tissues.

- Therefore, polynucleotides and polypeptides of the invention are useful as
30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, cancer, particularly of endothelial tissues. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system,

- 5 expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., endothelial, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in
- 10 healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 108 as residues: Thr-52 to Gly-57. Polynucleotides encoding said polypeptides are also provided.

- Expression within embryonic tissue and other cellular sources marked by
- 15 proliferating cells indicates that the protein product of this gene may play a role in the regulation of cellular division and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Additionally, the expression
 - 20 in hematopoietic cells and tissues indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic
 - 25 development also relies on decisions involving cell differentiation and/or apoptosis in pattern formation.

- Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain
- 30 neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have

applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types

- 5 of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new
- 10 insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or
- 15 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

- 20 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 929 of SEQ ID NO:16, b is an integer of 15 to 943, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where b is greater than or equal to a + 14.
- 25

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by

- 30 the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

RFLSVXPQXEVPFLLHPCVCXFXXGGHPSLLPDPCRAVGGGWEAPRCCLEALC
QLGCKAEEIVSVSESSAQRCWYLLRGRKAGGRGPASPVLFALMRLESLCH
LCLACLFRLPATRTVYCMNEAEIVDVALGILIESRKQXKACEQPALAGADNP
EHSPPCSVSPHTSSGSSSEEEDSGKQALXPGLSPQRPGGSSACRSPEEEE

- 5 EEDVLKYVREIFFS (SEQ ID NO: 204) . Polynucleotides encoding these polypeptides are also provided. Polynucleotides of the invention do not consist of the nucleic acid sequences shown as GeneSeq Accession Nos: V59595 and V59744, which are hereby incorporated herein by reference.

This gene is expressed primarily in a variety of immune cell types, including
10 stromal cells, dendritic cells, leukocytes, activated T-cells, macrophages, monocytes, neutrophils and to a lesser extent in a variety of other adult and fetal tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
15 not limited to, cancer and other proliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain
20 tissues or cell types (e.g., immune, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution in immune cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Expression of this gene product in fetal
30 tissue and various hematopoietic cancers indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all

hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

- 5 Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, 10 such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, scleroderma and tissues. Moreover, the protein may represent a secreted 15 factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, 20 raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly 25 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or 30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1489 of SEQ ID NO:17, b is an

integer of 15 to 1503, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

5 When tested against Jurkat T-cell lines, supernatants removed from cells containing this gene activated the NF- κ B (Nuclear Factor κ B) pathway. Thus, it is likely that this gene activates T-cells through the NF- κ B signal transduction pathway. NF- κ B is a transcription factor activated by a wide variety of agents, leading to cell activation, differentiation, or apoptosis. Reporter constructs utilizing the NF- κ B promoter element are used to screen supernatants for such activity. Preferred polypeptides of the invention comprise the following amino acid sequence: VPGWPRACSPCQADSPRAHPPKLRGILRWAPVPLXCAALCPPLDSG MSMAACPEAPEPSFLREVPSSPASTQWHRPCNFRQVEANPRKEPKNLVWRD VSLGQXSRTPRGSGLELVRVCAGGMQRDKTVVEERVGEERERERERESLGG

10 15 15 AGKHGEMRCVYVRESVGAPGRAGGGNGVNSVGCRTVHSGSXPPPSAGV S (SEQ ID NO: 205). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in parts of the brain such as cerebellum and frontal lobe.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous, or wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in cerebellum and frontal lobe indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, prevention and/or treatment of neurodegenerative disease states and behavioural disorders, or inflammatory conditions. Representative uses are described

5 in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms,

10 hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to

20 identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

25 related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1498 of SEQ ID NO:18, b is an

integer of 15 to 1512, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

5 In a specific embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

TRPGKELNLVFGLQLSMARIGSTVNMMNLMGWLYSKIEALLGSAGHTTLGITL

10 MIGGITCILSICALALAYLDQRAERILHKEQGKTGEVIKLTDVKDFSLPLWLIF
IICVCYYVAVFPFIGLGKVFFTEKFGFSSQAASAINSVVYVISAPMSPVFGLLV
DKTGKNIIWVLCA (SEQ ID NO: 206) . Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is believed to reside on chromosome

15 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

This gene is expressed primarily in fetal tissue, and to a lesser extent in a variety of adult human tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, fetal abnormalities, particularly developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For 20 a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developing, or cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having 25 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 111 as residues: Lys-30 to Thr-35. Polynucleotides encoding said polypeptides are also provided.

- The tissue distribution in fetal tissue indicates that polynucleotides and
- 5 polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.
- 10 Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Expression within embryonic tissue and other cellular sources marked by proliferating cells
- 15 indicates that this protein may play a role in the regulation of cellular division. Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating,
- 20 detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in
- 25 proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Additionally, the expression in hematopoietic cells and tissues indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene is useful in the treatment of lymphoproliferative disorders,
- 30 and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Furthermore, the protein

may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or

- 5 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

- 10 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1641 of SEQ ID NO:19, b is an integer of 15 to 1655, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:19, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

The translation product of this gene shares sequence homology with human histiocyte-secreted factor (HSF) which is a novel cytokine that shows in vivo

- 20 antitumour activity without the cytotoxicity associated with tumour necrosis factor. Furthermore, The translation product of this gene also shares sequence homology with the human endogenous virus S71 gag polyprotein, the sequence of which is believed to represent a transformation locus for several cancers (See Genebank Accession No. pir|A46312|A46312; all references available through this accession are hereby incorporated by reference herein). Similarly, The translation product of this gene also shares homology with B219, a sequence that is expressed in at least four isoforms in very primitive hematopoietic cell populations which may represent a novel hemopoietin receptor (See, e.g., Cioffi, et al. Nat. Med. 2:585-589 (1996), which is hereby incorporated by reference herein). In a preferred embodiment
- 25 30 polypeptides of the invention comprise the following amino acid sequence:

CKDLCRVYLLTLSPLLSYDPATSHSPRNTQ (SEQ ID NO: 207) . Also preferred are the polynucleotides encoding these polypeptides.

This gene is expressed primarily in tonsil, and colon, and to a lesser extent in a wide variety of human tissues.

- 5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hematopoietic, and gastrointestinal disorders, particularly tumors of the colon and tonsil. Similarly, polypeptides and antibodies directed to
- 10 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic, digestive and immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, gastrointestinal, or
- 15 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.
- 20 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 112 as residues: Met-1 to Cys-6. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in tonsil and colon, combined with the homology to human histiocyte growth factor, the human endogenous viral protein, and B219

25 strongly indicate that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment and/or prevention, of a variety of hematopoietic and immune system disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19,

30 20, and 27, and elsewhere herein. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation

of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g., by boosting immune responses).

- 5 Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, 10 such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, scleroderma and tissues. Moreover, the protein may represent a secreted 15 factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, 20 raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly 25 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or 30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2511 of SEQ ID NO:20, b is an

integer of 15 to 2525, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

5 The gene encoding the disclosed cDNA is believed to reside on chromosome 7. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 7.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by 10 the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

IICECWEEEQSCRLKITQPREICRMDFLVFLFYLASVLMGLVLICVCSKTHSLKGLARGGAQIFSCIPECLQRAXHGLLHYLFHTRNHTFIVLHLVLQGMVYTYTWEVFGYCQELELSLHYLLPYLLLGVNLFFFTCGTNPGIITKANEELLFLH
15 VYEFDEVMPKNCSTCDLRKPARSKHCSVCNWCVHRFDHHCVWVNNCIGAWNIRYFLIYVLTLTASAATVAIVSTTFLVHLVVMSDLYQETYIDDLGHLHVMDTVFLIQYLFLTFPRIVFMLGFVVVLSFLLGGYLLFVLYLAATNQTTTNEWYRGDWAWCQRCPLVAWPPSAEPQVHRNIHSHGLRSNLQEIFLPAPCHERKKQE(SEQ ID NO: 208) . Polynucleotides encoding these polypeptides are also provided.

20 This gene is expressed primarily in colon and brain and to some extent in all tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are 25 not limited to, neurological and digestive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and digestive system, expression of this gene at significantly higher or lower levels is 30 routinely detected in certain tissues or cell types (e.g., neurological, gastrointestinal, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,

synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly,
- 10 the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive
- 15 compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

- Potentially, this gene product is involved in synapse formation,
- 20 neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Alternatively, expression of this gene in colon may indicate a role in the detection, prevention and/or treatment of colon disorders such as colon cancer, Crohn's Disease, ulcers, and digestive tract disorders in general. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to
- 25 isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly
- 30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

- 5 formula of a-b, where a is any integer between 1 to 1382 of SEQ ID NO:21, b is an integer of 15 to 1396, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 12

- 10 When tested against Reh cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activation site) pathway. Thus, it is likely that this gene activates B-cells through the Jaks-STAT signal transduction pathway. GAS is a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.
- 15 Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. This gene maps to chromosome 7, and therefore, is used as a marker in linkage analysis for chromosome 7.
- 20 This gene is expressed primarily in brain, and in the developing embryo. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, behavioral, immune, and developmental disorders.
- 25 Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and developmental systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural,
- 30 developing, immune, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue

or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 114 as residues: Lys-60 to Asn-67. Polynucleotides encoding said polypeptides are also provided.
- The tissue distribution in brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions.
- 10 Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, 15 trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the tissue distribution in developing 20 embryo indicates that the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Alternatively, the biological activity within B-cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of 25 immune system disorders. Activation of genes within B-cells indicates a role for this protein in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of 30 cancer (e.g., by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne,

5 and psoriasis. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions; in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

10 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
15 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1055 of SEQ ID NO:22, b is an integer of 15 to 1069, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where b is greater than or equal to a + 14.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

The gene encoding the disclosed cDNA is believed to reside on chromosome 6. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 6.

25 In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

LLSFKIRGLRTEDAGWAQSSSGGLCVRGDAFWMPSSSGLGSPSRPPSSLCL

30 LLLLLPPAALALLFFLDFFPPRAAVSPFLPDHCSARQPRVWRRETLNRSASGL

GCWARSTEQGAVGVATGTVLDI SLPASCLSLWPPGPSGGI (SEQ ID NO: 209)

. Polynucleotides encoding these polypeptides are also provided.

In a specific embodiment polypeptides of the invention comprise the following amino acid sequence:

- 5 QLGLCLTSASLPPASRCGHQAPLGASDLSAHSAPGFSDSYFTMSCQSSLSRA
EILQCPLVPSVSPPTHLPQGRANKSSRASLPLPQTHWCLFPSARGWRRGIQSG
LPPGGSCTSPRSPPQTLHQHITLVNHNTSYWQSPST (SEQ ID NO: 210) ,
HQPPCLLPLAVATRPLWGHLTCLPIILHLVSVTLTSPCLANQAFQGQRQSYNAL
WCPLFLLLPTSPKGEGQTNHPEPACPCFPKLTGVFSLQHVVGAEFSQVFLLVD
10 PVPVLHDLLKLFTSTSHLLIIIPHIGKAPAPDSLL EELSLSLATHCKVAVARFT
(SEQ ID NO: 211) . Also preferred are the polynucleotides encoding these polypeptides. Polynucleotides of the invention do not consist of the nucleic acid sequence shown as GeneSeq Accession No. X04377, which is hereby incorporated herein by reference.

- 15 This gene is expressed primarily in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, behavioral and neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 115 as residues: Pro-2 to Gly-7, Ser-10 to Ser-16,

Pro-52 to Val-62, Arg-64 to Ser-73. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of

- 5 neurodegenerative disease states, behavioral disorders, or inflammatory conditions.
- Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome,
- 10 meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep
- 15 patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity,

- 20 to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly
- 25 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
 - 30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1644 of SEQ ID NO:23, b is an

integer of 15 to 1658, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 14

- 5 The translation product of this gene was shown to have homology to the lysosomal mannosidase alpha-B protein (See Genebank Accession No. P34098) which is thought to be important in protein metabolism. One embodiment of this gene comprises polypeptides of the following amino acid sequence:
- MAAEGSRFSSQSPGLVDRQGPKCDPSRLVSPWGRHGLRILQIGHHHGRDGQH
- 10 EATHHLLRVLRAPRVGKADEGAVDSDPSTPLQLKHEAAHAEDHAQQVHVVR
RRVVQGRVTFARRGLVPQHFVRPPWVRHTVSGHSESKARSRLFRCRNRSFRR
AS (SEQ ID NO: 212) , and/or
- 15 RLVSPWGRHGLRILQIGHHHGRDGQHEATHHLL RVLRA (SEQ ID NO: 213) .
An additional embodiment is the polynucleotides encoding these polypeptides. This gene maps to chromosome 19, and therefore, is used as a marker in linkage analysis for chromosome 19.

This gene is expressed primarily in brain, placenta, fetal liver, and to a lesser extent in most tissues.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, reproductive, and hepatic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For 20 a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, hepatic, or cancerous and wounded tissues) or bodily fluids (e.g., bile, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a 25 disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 116 as residues: Asn-34 to Lys-42, Leu-60 to Trp-70. Polynucleotides encoding said polypeptides are also provided.

- The tissue distribution predominantly in brain indicates a role in the
- 5 detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder.
- Alternatively, the tissue distribution in liver indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of
- 10 liver disorders and cancers (e.g., hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma.
- 15 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
- 20 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1063 of SEQ ID NO:24, b is an integer of 15 to 1077, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:24, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 15

This gene is expressed primarily in spinal cord, Merkel cells, and adipose tissues.

- Therefore, polynucleotides and polypeptides of the invention are useful as
- 30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, disorders of the nervous and immune systems, particularly those disorders relating to the CNS involving lipid metabolism disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For

- 5 a number of disorders of the above tissues or cells, particularly of the nervous and immune systems and adipose tissue, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, immune, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an
- 10 individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in spinal cord, Merkel cells and adipose tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the

- 15 treatment and/or diagnosis of diseases the nervous systems, such as spinal cord injury, neurodegenerative diseases, muscular dystrophy or obesity. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly
- 20 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
 - 25 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1191 of SEQ ID NO:25, b is an integer of 15 to 1205, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where b is greater than or equal to a + 14.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 16

The translation product of this gene shares sequence homology with the human uncoupling protein-2 which is thought to be important in energy metabolism, obesity, and the predisposition of hyperinsulinemia (See Genebank Accesion No. gi|1857278). Recently, another group published on this gene, designating it brain mitochondrial carrier protein-1 (BCMP1) (J Biol Chem 1998 Dec 18;273(51):34611-5). One embodiment of this gene comprises polypeptides of the following amino acid sequence: PTDVLKIRMQAQ (SEQ ID NO: 214), and/or TYEQLKR (SEQ ID NO: 215). An additional embodiment is the polynucleotides encoding these polypeptides. This gene maps to the X chromosome, and therefore, is used as a marker in linkage analysis for the X chromosome.

This gene is expressed primarily in manic depression brain tissue, epileptic frontal cortex, human erythroleukemia cell line, T-helper cells, and to a lesser extent in endothelial and amygdala cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the central nervous system or hematopoietic/immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system or hematopoietic/immune systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, hemolymphoid, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 118 as residues: Ser-34 to Thr-39, Gln-198 to Leu-205. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in neural tissues combined with the homology to the human uncoupling protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and/or treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's

- 5 Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of
- 10 developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Alternatively, given the homology to uncoupling proteins, the gene and/or its translation product may also be used in the diagnosis, treatment, and/or prevention of thermogenesis disorders such as obesity, cachexia, and hyperinsulinemia. Uncoupling proteins dissipate the proton gradient created from the oxidation of fuels by the electron transport chain, thus releasing stored energy as heat. Dysfunction of thermogenesis can induce disorders such as obesity and cachexia. It is thought that obesity may result from decreased thermogenesis in humans. Alternatively, cachexia is a metabolic state in which energy expenditure exceeds food intake, for example in anorexia nervosa. Uncoupling proteins is useful for the treatment and/or prevention of diseases and/or disorders involved with aberrant metabolic and thermogenic pathways. The following method provides for the determination of respiration uncoupling activity of the polypeptides of the present invention, including fragments and variants of the full length proteins.

Briefly, yeast are transfected with an expression vector expressing polypeptide of the present invention as previously described by Bouillaud et al., EMBO J., 13:1990 (1994) (incorporated by reference herein in its entirety). Rates of growth in liquid medium of transformed yeast are measured in the presence of galactose, which induces expression, as described in International Publication No. WO 98/31396

(incorporated by reference herein in its entirety). Instantaneous generation times are compared between the polypeptide of the present invention and appropriate controls. An in vivo decrease of membrane potential associated with uncoupling of respiration is analyzed by flow cytometry of yeast labeled with the potential sensitive probe

- 5 DiOC₆ (3) (3,3'-dihexyloxacarbocyanine iodine, Molecular Probes, Eugene, OR). The ability of a polypeptide of the present invention to influence mitochondrial activity and uncouple respiration is thus determined.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are 10 related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general 15 formula of a-b, where a is any integer between 1 to 1660 of SEQ ID NO:26, b is an integer of 15 to 1674, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

- 20 The translation product of this gene shares sequence homology with 55 kD deglycosylated zona pellucida protein which is known to be important in egg fertilization (See Genebank Accession No.R39356). Preferred polypeptides of the invention comprise the following amino acid sequence:
- 25 RPRPSASSLARSASLLPAAHGXGVGGAGGGSSXLSRYQQQLQNEEESGEPEQ
AAGDAPPPYSSISAESAHXFDYKDESGFPKPPSYNVATLPSYDEAERTKAEA
TIPLVPGRDEDVFGRDDFDDADQLRIGNDGIF (SEQ ID NO: 216),
RYQQLQNEEESGEPEQAAGD (SEQ ID NO: 217), and/or
PGRDEDVFGRDDFDDADQLRIG (SEQ ID NO: 218). Polynucleotides encoding 30 these polypeptides are also provided.

Preferred polypeptide fragments of the invention comprise the following amino acid sequence: MLTFFMAFLFNWIGFFLSFCLTTSAAGRYG
AISGFGLSLIKWILIVRFSTYFPGYFDGQY
WLWWVFLVLGFLFLRGFINYAKVRKMPET FSNLPRTRVLFIY (SEQ ID NO: 5 219). Polynucleotides encoding these polypeptides are also provided.

Preferred polypeptide variants of the invention comprise the following amino acid sequence:

MKKSLENLNRLQVMILLHLTA AFLQRAQH邢D YKDESGFPKPPSYNVATTLPS
YDEAERTKAEATIPL VPGRDEDVGRDDFDDADQLRIGNDGIFMLTFFMAFLF
10 NWIGFFLSFCLTTSAAGRYGAISGFGLSLIKWILIVRFSTYFPGYFDGQYWLW
WVFLVLGFLFLRGFINYAKVR KMPETFSNLPRTRVLFIY (SEQ ID NO: 220),
MLLHLTA AFLQRAQFSTYFPGYFDGQYWLWVFLVLGFLFLRGFINYAKV
RKMPETFSN LPRTRVLFIY (SEQ ID NO: 221), MLTFFMAFLFNWIGFFLSFCLT
TSAAGRYGAISGFGLSLIKWILIVRFSTYFPAFMNSLSRSKRTPAGSESRCRTQ
15 RNNHLL (SEQ ID NO: 222), and/or
MKKSLENLNRLQVMILLHLTA AFLQRAHXIL TTRMSLGFQSPHLM (SEQ ID
NO: 223) . Polynucleotides encoding these polypeptides are also provided.

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other immune and hematopoietic cells and JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the 20 Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in adult kidney, colon adenocarcinoma, and fetal brain, and to a lesser extent, ubiquitously expression in many tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as 30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, disorders of kidney, colon cancers, and central nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the

- 5 renal and neural systems, and cancers, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., renal, neural, urogenital, or cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression
10 level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 119 as residues: Cys-15 to Gly-36. Polynucleotides encoding said polypeptides are also provided.

- 15 The tissue distribution adult kidney, colon adenocarcinoma, and fetal brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of kidney diseases, colon cancers, and disorders of the central nervous system. Additionally, the homology to the zona pellucida protein indicates that the gene product is used for male contraceptive development, and
20 infertility diagnosis etc. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
25 related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
30 formula of a-b, where a is any integer between 1 to 1951 of SEQ ID NO:27, b is an

integer of 15 to 1965, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

- 5 The translation product of this gene shares sequence homology with the chicken transferrin receptor in addition to a human prostate-specific protein homolog (See Genbank Accession Nos.pir|JH0570|JH0570 and gi|2565338 (AF026380), respectively). This gene also shares significant homology with both the murine and the rat hematopoietic lineage switch 2 proteins (See Genbank Accession Nos. 10 g3169729 and g3851632, respectively), which are induced during an erythroid to myeloid lineage switch.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MTVMDPKQMNVAAGWAVVSYVVADMEEML PRS (SEQ ID NO: 224). Polynucleotides encoding these polypeptides are also provided.

- 15 This gene is expressed primarily in fetal tissues, such as liver/spleen and brain. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pre-natal disorders, anomalies, deficiencies. Similarly, polypeptides 20 and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developing fetus, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developing, cancerous and wounded tissues) or bodily fluids 25 (e.g., amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 120 as residues: Arg-31 to Lys-37, Lys-58 to Glu-65,

Asp-157 to Gly-168, Ile-219 to Gly-225, Ala-260 to Ser-268, Thr-276 to Glu-282.

Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of pre-natal

- 5 disorders, anomalies and deficiencies. The homology to the hematopoietic lineage switch 2 proteins indicates that The translation product of this gene is useful for the detection and/or treatment of immune system disorders. In addition, the homology to the transferrin receptor indicates that the translation product of the present invention may have utility in the regulation of iron metabolism as well as the
10 numerous genes under the stringent control of physiologic iron levels. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

- 15 related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
20 formula of a-b, where a is any integer between 1 to 1849 of SEQ ID NO:28, b is an integer of 15 to 1863, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

- 25 In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:
PRVRSREPVAGAPGCGTAGPPAMATLWGGLRLGSLLSCLALSVLLAHC
30 QTPPSDCLHVVEPMPVRGPDVAYCLRCECKYEERSSVTIKVTIIYLSILGLL
LYMVYLTLVEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDLARSRSRANV

LNKVEYAQQRWKLQVQEQRKSVFDRHVVLS (SEQ ID NO: 225).

Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 72 - 88 of the amino acid sequence referenced in

- 5 Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 89 to 167 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia membrane proteins.

A preferred polypeptide variant of the invention comprise the following amino

- 10 acid sequence:

MATLWGGLLRGSLLSCLALSVLLAHCQTTPRISRMSDVNVSAALPIKKNS
GHIYNKNISQKDCDCLHVVEPMPVRGPDVAYCLRCECKYEERSSVTIKVTIII
YLSILGLLLLYMVYLTLEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDVLA
RSRSRANVLNKVEYGTAALEASSPRAAKSLSLTGMLSSANWGIEFKVTRKKQ

- 15 ADNWKGTDWVLLGFILIPC (SEQ ID NO: 226). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in infant brain tissue, and to a lesser extent in other cell types and tissues.

- Therefore, polynucleotides and polypeptides of the invention are useful as
20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and neurodegenerative diseases of the brain and nervous system, such as depression, schizophrenia, Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, mania, dementia, paranoia, addictive behavior, sleep
25 disorders, epilepsy, transmissible spongiform encephalopathy (TSE), Creutzfeldt-Jakob disease (CJD). Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher
30 or lower levels is routinely detected in certain tissues or cell types (e.g., neural, developmental, or cancerous and wounded tissues) or bodily fluids (e.g., amniotic

fluid, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 121 as residues: Gln-110 to Pro-120, Val-152 to Val-159. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in infant brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis
10 of developmental, degenerative and behavioral conditions of the brain and nervous system. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of schizophrenia, Alzheimer's Disease, Parkinson's
15 Disease, Huntington's Disease, Tourette Syndrome, transmissible spongiform encephalopathy (TSE), Creutzfeldt-Jakob disease (CJD), mania, depression, dementia, paranoia, addictive behavior, obsessive-compulsive disorder and sleep disorders. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to
20 identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
25 related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
30 formula of a-b, where a is any integer between 1 to 1612 of SEQ ID NO:29, b is an

integer of 15 to 1626, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

5 The translation product of this gene shares sequence homology with a recently published gene Dysferlin, which is thought to be a skeletal muscle gene that is mutated in Miyoshi myopathy and limb girdle muscular dystrophy (See Genbank Accession No. g3600028, and Nat. Genet. 20 (1), 31-36 (1998)).

This gene is expressed primarily in fetal liver, fetal heart tissue, and T-cells.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immunodeficiency, tumor necrosis, lymphomas, auto-immunities, cancer, inflammation, anemias (leukemia) and liver disorders, vascular disorders, and
15 cancers (e.g., hepatoblastoma, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the liver and
20 immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., hepatic, developmental, vascular, or cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, bile, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard
25 gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immunities, immunodeficiencies
30 (e.g., AIDS), immuno-suppressive conditions (transplantation) and hematopoietic disorders. In addition this gene product is applicable in conditions of general

microbial infection, inflammation or cancer. Expression in liver may suggest a role for this gene product in the treatment and detection of liver disorders and cancers (e.g., hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). Alternatively,

- 5 the tissue distribution in fetal heart tissue indicates that the protein product of this gene is useful for the diagnosis and treatment of conditions and pathologies of the cardiovascular system, such as heart disease, restenosis, atherosclerosis, stroke, angina, thrombosis, and wound healing. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets
- 10 for the above listed tissues. Additionally, the homology to the dysferlin gene indicates that polynucleotides and polypeptides corresponding to this gene are useful for diseases related to degenerative myopathies that are characterized by the weakness and atrophy of muscles without neural degradation; such as Duchenne and Becker's muscular dystropies. Protein, as well as, antibodies directed against the protein may
- 15 show utility as a tumor marker and/or immunotherapy targets for the above listed tissues

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of

- 20 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 591 of SEQ ID NO:30, b is an
- 25 integer of 15 to 605, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

- This gene is expressed primarily in haemopoietic cells and tumor cells, such
- 30 as pancreatic tumor tissue, and to a lesser extent in bladder cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, haemopoietic disorders, diseases of the renal and pancreatic systems,

- 5 and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the haemopoietic, pancreatic, and renal systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or
10 cell types (e.g., pancreas, renal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 15 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the renal, pancreatic and haemopoietic systems, and cancers thereof. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 20 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
25 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 917 of SEQ ID NO:31, b is an integer of 15 to 931, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where b is greater than or equal to a + 14.

This gene is expressed primarily in liver tissue, cancer cells and fetal lung tissue, and to a lesser extent in dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

- 5 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, metabolic disorders, diseases of developing systems and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the
10 fetus, metabolic systems and cancers, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developing, metabolic, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression
15 level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 124 as residues: His-44 to Gly-49. Polynucleotides encoding said polypeptides are also provided.

- 20 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the fetus, metabolic systems and cancers. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice,
25 hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma. Protein, as well as, antibodies directed against the protein may show utility as
30 a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

- 5 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1393 of SEQ ID NO:32, b is an integer of 15 to 1407, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where b is greater than or equal to a + 14.
- 10

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

This gene is expressed primarily in central nervous system tissues and cancers, such as endometrial tumors, and to a lesser extent in other tissues and organs.

- 15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the CNS and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological
- 20 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and cancerous tissues, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.
- 25

- Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 125 as residues: Tyr-16 to Ser-22, Asp-209 to Leu-215. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of diseases of the central nervous system, as well as cancers of tissues where expression of this gene has been observed, such as in 5 endometrial tumors. The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive 10 disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein 15 may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of 20 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1512 of SEQ ID NO:33, b is an 25 integer of 15 to 1526, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:33, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

The translation product of this gene shares sequence homology with low-density lipoprotein receptor (See Genbank Accession No. >dbj|BAA24580.1), which 30 is thought to be important in the pathogenesis of atherosclerosis and other disorders.

The translation product of this gene also shares sequence homology with a rat homolog of the human CD94 (See Genbank Accession No. gb|AAC10220.1).

This gene is expressed primarily in macrophages, eosinophils, neutrophil and other cells of the haemopoietic and immune system.

- 5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the immune and haemopoietic systems and diseases of the endothelial and vascular system. Similarly, polypeptides and antibodies directed to
- 10 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, haemopoietic and vascular system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily
- 15 fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 126 as residues: Lys-9 to Ala-17, Met-55 to Leu-61, Tyr-105 to Cys-127, Asp-132 to Lys-141, Ser-165 to Tyr-172, Pro-178 to Met-186, His-222 to Gln-227. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution and homology to LDL receptor and rat CD94 homolog indicates that polynucleotides and polypeptides corresponding to this gene are useful

25 for the treatment and/or diagnosis of disorders of the immune, haemopoietic and vascular systems. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of hematopoietic disorders. This gene product is primarily expressed in hematopoietic cells and tissues, suggesting that it plays a role in the survival, proliferation, and/or

30 differentiation of hematopoietic lineages. Expression of this gene product in eosinophils and macrophage also strongly indicates a role for this protein in immune

function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly
- 5 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
- 10 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1723 of SEQ ID NO:34, b is an integer of 15 to 1737, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where b is greater than or equal to a + 14.

15 **FEATURES OF PROTEIN ENCODED BY GENE NO: 25**

A preferred polypeptide fragment of the invention comprises the following amino acid sequence:

- MAAAGRLPSSWALFSPLLAGLALLGVGPVPARALHNVTAEFGAEAWGTLA
AFGDLNSDKQTDLFVLRERNDLIVFLADQNAPYFKPKVKVSFKNHSALITSVV
20 PGDYDGDSQMDVLLTYLPKNYAKSELGAVIFWGQNQTLDPNNMTILNRTFQ
DEPLIMDFNGDLIPDIFGITNESNQPQILLGGNLSWHPALTTSKMRIPHSHAFI
DLTEDFTADLFLLTTLNATTSTFQFEIWEENLDGNFSVSTILEKPQNMMVGQSA
FADFDDGDGHMDHLLPGCEDKNCQKSTIYLVRSGMKQWVPVLQDFSNKGTL
WGFVPFVDEQQPTEIPIPI TLHIGDYNMDGYPDALVILKNTSGSNQQAFLLENV
25 PCNNASCEEARRMFKVYWELTDLNQIKDAMVATFFDIYEDGILDIVVLSKGY
TKNDFAIHTLKNNFEAADAYFVKVIVLGLCS NDCPRR (SEQ ID NO: 227).

Polynucleotides encoding these polypeptides are also provided.

- When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter
- 30 element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other immune and hematopoietic cells and tissue cell types, through the JAK-STAT

signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the

- 5 GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in infant brain and placental tissues, and to a lesser extent in several other tissues including cancers.

Therefore, polynucleotides and polypeptides of the invention are useful as

- 10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, brain disorders and diseases of developing systems and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell
15 type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and fetal systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, developing, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken
20 from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 127 as residues: Leu-56 to Thr-62, Gln-80 to Pro-87, 25 Gly-106 to Gln-113, Pro-122 to Lys-127, Gln-138 to Asn-146. Polynucleotides encoding said polypeptides are also provided.

- The tissue distribution in neural tissues and developing tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of the central nervous system, disorders of 30 developing systems, and cancers. The tissue distribution in infant brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful

for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered

- 5 behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product is produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta also indicates that this gene product is produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 25 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

formula of a-b, where a is any integer between 1 to 2228 of SEQ ID NO:35, b is an integer of 15 to 2242, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:35, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 26

Preferred polypeptides of the invention comprise the following amino acid sequence:

MTKREDGGYTFTATPEDFPKKHKAPVIDIGIANTGKFIMTASSDTTVLIWSLK
GQLVSTINTNQMNNTHAAVSPCGRFVASCFTPDVKVWEVCFGKKGEFQEVEV

10 VRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTXVEYKKQDPY
LLKTGRFEAAAGAXPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH
GECIANLSFDITGRFLASCGDRAVRLFHNTPGHGRAMVEEMQGHLKRASNEST
RQRLQQQLTQAQETLKGALKK (SEQ ID NO: 228). Polynucleotides encoding such polypeptides are also provided.

15 The gene encoding the disclosed cDNA is thought to reside on chromosome 7. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 7.

When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates 20 myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS 25 element, can be used to indicate proteins involved in the proliferation and differentiation of cells. Recently, another group published this gene, naming it WS beta-transducin repeats protein (See Human Genetics 103 (5), 590-599 (1998); which is hereby incorporated herein by reference), in which it was suggested that the protein is involved in William's Disease.

30 The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 12 - 28 of the amino acid sequence referenced in

Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by

- 5 the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

VIRHEGSTNMELSQMSXLMGLSVLLGLLALMATAAVXRGWLAGEERSGRP
ACQKANGFPPDKSSGSKKQKQYQRIRKEKPQQHNFTHRLAALALKSHSGNIS
CMDFSSNGKYLATCADDRTIRIWSTKDFLQREHRSMRANVELDHATLVRFSP

- 10 DCRAFIVVLANGDTLRFVKMTKREDGGYTFTATPEDFPKKHKAPVIDIGIAN
TGK

FIMTASSDDTTVLIWSLKGQVLSTINTNQMNNTHAAVSPCGRFVASCGFTPDK
VWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWK
LWDTXVEYKKKQDPYLLKTGRFEEAAGAXPCRLALSPNAQVLALASGSSIHL

- 15 YNTRRGEKEECFERVHGECIANLSFDITGRFLASCGDRAVRLFHNTPGHGRAM
VEEMQGHLKRASNESTRQRLQQQLTQAQETLKGALKK (SEQ ID NO: 229).
Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in testes, synovial sarcoma and fetal tissues, and to a lesser extent in several other tissues.

- 20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the reproductive and developing systems and cancers.

- Similarly, polypeptides and antibodies directed to these polypeptides are useful in
25 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and developing systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, testicular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, seminal fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in testes tissue, synovial sarcoma, and fetal tissues, indicates that polynucleotides and polypeptides corresponding to this gene are useful

- 5 for the treatment and/or diagnosis of disorders of the reproductive and developing systems, and cancers. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g., endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment
- 10 of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that is expressed, particularly at low levels, in other tissues
- 15 of the body. Therefore, this gene product is expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications.

Furthermore, the tissue distribution indicates that polynucleotides and

- 20 polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Additionally, the expression in hematopoietic cells and tissues indicates that this protein may play a role in the
- 25 proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene is useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in
- 30 pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. This protein is useful for

the treatment, detection, and/or prevention of William's Disease. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,

- 5 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:36 and may have been publicly available prior to conception of

- 10 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2221 of SEQ ID NO:36, b is an
15 integer of 15 to 2235, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:36, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

Preferred polypeptides of the invention comprise the following amino acid

- 20 sequence positions 1-363, 2-363, 4-363, 5-363, 30-363, 31-363, 32-363, 75-363, 76-363 and 82-363 of the following amino acid sequence:

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTLFLILGTCTL
FFAFECRYLAQLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIE
MEIEATNGAVPQGQRPPPRIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDN
25 CVERFDHHCPWVGNCVGKRNYRYFYLFILSLSLLTIYVFANIVYVALKSLKI
GFLETLKETPGTVLEVLCFTLWSVVGLTGFHTFLVALNQTTNEDIKGWSWTG
KNRVQNPYSHGNIVKCCEVLCGPLPPSVLDRRGILPLEESGSRPPSTQETSSS
LLPQSPAPTEL NSNEMPEDSSTPEEMPPPEPPQEEAAEK (SEQ ID NO:
230). Polynucleotides encoding such polypeptides are also provided.

- 30 A preferred polypeptide variant of the invention comprises the following amino acid sequence: MLFLFSMATLLRTSFSDPGVIPRALPDEAA

FIEMEIEATNGAVPQGQRPPPRIKNFQINNQIVKLKYCYTCKIFRPPRASHCSIC
DNCVE RFDHHCPWVGNCVGKRNRYFYLFILSLLTIYVFAFNIVYVALK
SLKIGFLETLKGNS WNCRSRSPHLLLTYLVRRGTDWISYFPRGSQ PDNQ (SEQ
ID NO: 231). Polynucleotides encoding these polypeptides are also provided.

- 5 This gene is expressed primarily in ovarian and endometrial tumors, fetal liver, spleen and brain tissues, and to a lesser extent in several other tissues and organs.
- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
- 10 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the developing systems, and cancers of the female reproductive system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above
- 15 tissues or cells, particularly of the developing, female reproductive and fetal systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, developing, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 129 as residues: Pro-44 to Lys-54, Cys-88 to His-95, Val-103 to Tyr-108, Gln-181 to Ser-190, Thr-192 to Ile-206, Glu-233 to Ser-245, Ser-252 to Ala-286. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in developing systems indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of disorders of developing and fetal systems, and cancers. Furthermore, the tissue distribution in ovarian and endometrial tumor tissues indicates that the translation product of this gene is useful for the detection, diagnosis, and/or treatment of cancers of the female reproductive system. Accordingly, preferred are antibodies

which specifically bind a portion of The translation product of this gene. Also provided is a kit for detecting these tumors. Such a kit comprises in one embodiment an antibody specific for The translation product of this gene bound to a solid support. Also provided is a method of detecting these tumors in an individual which

- 5 comprises a step of contacting an antibody specific for The translation product of this gene to a bodily fluid from the individual, preferably serum, and ascertaining whether antibody binds to an antigen found in the bodily fluid. Preferably the antibody is bound to a solid support and the bodily fluid is serum. The above embodiments, as well as other treatments and diagnostic tests (kits and methods), are
- 10 more particularly described elsewhere herein.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

- 15 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2957 of SEQ ID NO:37, b is an integer of 15 to 2971, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:37, and where b is greater than or equal to a + 14.
- 20

FEATURES OF PROTEIN ENCODED BY GENE NO: 28

This gene is expressed primarily in normal and cancerous colon tissue, macrophages, endothelial cells and placental tissue, and to a lesser extent in several other tissues and organs.

- 25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, colon cancer and gastrointestinal disorders, immune disorders, vascular diseases and disorders of developing systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for
- 30

differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, vascular and developing systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, gastrointestinal, developmental,

- 5 vascular, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 10 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 130 as residues: Thr-27 to Ser-33. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in macrophage, endothelial and placental tissues, and normal and cancerous colon tissues, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of immune, gastrointestinal and vascular disorders and diseases. Expression of this gene product in colon tissue indicates involvement in digestion, processing, and elimination of food, as well as a potential role for this gene as a diagnostic marker or causative agent in the development of colon cancer, and cancer in general. Accordingly, preferred are antibodies which specifically bind a portion of the translation product of this gene. Also provided is a kit for detecting colon cancer. Such a kit comprises in one embodiment an antibody specific for the translation product of this gene bound to a solid support. Also provided is a method of detecting colon cancer in an individual which comprises a step of contacting an antibody specific for the translation product of this gene to a bodily fluid from the individual, preferably serum, and ascertaining whether antibody binds to an antigen found in the bodily fluid. Preferably the antibody is bound to a solid support and the bodily fluid is serum. The above embodiments, as well as other treatments and diagnostic tests (kits and methods), are more particularly described elsewhere herein. Alternatively, the tissue distribution in placental tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders

of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product is produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or

- 5 survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta and endothelial cells also indicates that this gene product is produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other
- 10 cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Additionally, expression of this gene product in macrophage also strongly indicates a role for this protein in immune function and immune surveillance. This gene product is involved in the regulation of cytokine
- 15 production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it is also used as

- 20 an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as,
- 25 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:38 and may have been publicly available prior to conception of

- 30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1149 of SEQ ID NO:38, b is an integer of 15 to 1163, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:38, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

The translation product of this gene shares homology with HNK-sulfotransferase, which directs glycan synthesis (see Genbank Accession no.

10 AF033827).

This gene is expressed primarily in activated T cells, osteoclastoma, and glioblastoma, and to a lesser extent in various other normal and transformed cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation, immune defects, cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hemopoietic systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative 25 to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 131 as residues: Pro-32 to Gly-48, Gln-63 to Thr-69, Pro-77 to Trp-84, Val-88 to Leu-94. Polynucleotides encoding said polypeptides are 30 also provided.

The tissue distribution in T-cells and various types of neoplasms indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, study and/or treatment of inflammatory and general immune defects, and various types of neoplasms. Expression of this gene product in T cells strongly

- 5 indicates a role for this protein in immune function and immune surveillance. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the gene or protein, as
10 well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility
15 in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the tissue distribution in various cancerous tissues indicates that the translation product of the gene is useful for the detection, diagnosis, and/or treatment of these cancers, as well as cancers of other tissues where expression has been observed. Protein, as well
20 as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of
25 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1918 of SEQ ID NO:39, b is an
30 integer of 15 to 1932, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

Preferred polypeptides of the invention comprise the following amino acid sequence:

- 5 LHECLPGSISYLPRTPWLCPPQHLSFSTFSPPWQPAMSPVPGTGGPPCGL
(SEQ ID NO: 232), and/or
MLPLLIICLPAIEGKNCRLCWPELSALIDYDLQILWVTGPPTELSQSIIHSLFLE
DNNFLKPWYLDRDHLEEETAKFFTQVHQAIKTLRDDKTVLLEIYTHKNLFT
ERLNKISDGLKEKGAPPLHECLPGSISYLPRTPWLCPPQHLSFSTFSPPWQP
10 AMSPVPGTGGPPCGL (SEQ ID NO: 233). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in infant brain, testes and activated T cells, and to a lesser extent in various other normal and transformed cell types.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, reproductive and inflammatory conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For 20 a number of disorders of the above tissues or cells, particularly of the neural, immune and male reproductive systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, immune, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken 25 from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 132 as residues: Gly-41 to Leu-46, Asp-67 to Thr-75, 30 Ile-114 to Ala-123. Polynucleotides encoding said polypeptides are also provided.

- The tissue distribution in infant brain tissue, testes tissue, and activated T-cells, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis, and/or treatment of neurological, reproductive and immune system disorders. Expression of this gene product in T-cells indicates a role
- 5 in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses).
- 10 Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it is also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease,
- 15 sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell type. Alternatively, the tissue distribution in testes tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions
- 20 concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis
- 25 of testicular cancer. The testes are also a site of active gene expression of transcripts that is expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product is expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications.
- 30 Furthermore, the tissue distribution in infant brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the

detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors,

- .5 including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

10 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
15 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 867 of SEQ ID NO:40, b is an integer of 15 to 881, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where b is greater than or equal to a + 14.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 31

The translation product of this gene shares sequence homology with some human and rodent melanoma and leukocyte specific antigens (see, for example, Genbank accession nos: gi|189384, gi|205898 and gi|180926). In addition, The

- 25 translation product of this gene shares sequence homology with Tetraspan protein (see, for example, Genbank accession number: GI 3152703). Therefore, it is likely that the polypeptide of this gene shares some biological functions, such as cell-to-cell signaling, adhesion, proliferation, and differentiation with Tetraspan.

The polypeptide of this gene has been determined to have two transmembrane domains at about amino acid position 52-68 and 197 - 213 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed

that the protein product of this gene shares structural features to type IIIa membrane proteins.

The transmembrane 4 superfamily (TM4SF) or tetraspan superfamily has at least 16 members (including CD9, CD20, CD37, CD53, CD63, CD81, CD82, A15,

- 5 CO-029, Sm23, RDS, Uro B, Uro A, SAS, Rom-1, PETA3, and YKK8), is the second biggest subfamily among CD antigen superfamily. and activation antigen of T- cells. All TM4SF member contains four putative transmémbrane domains, two extracellular loops, and two short cytoplasmic tails. They are variously expressed on Immature, early, mature, activated lymphocytes, monocytes, macrophages, granulocytes, platelets, eosinophils, basophils, certain leukemic and lymphoma cells, and a variety of other cells and tissues. CD9 cell surface protein is expressed by both hematopoietic and neural cells, and may play a role for CD9 in intercellular signaling in the immune and nervous system. CD63 is a 53-Kd lysosomal membrane glycoprotein that has been identified as a platelet activation molecule, which play important role in cell adhesion of platelets and endothelial cells. Increased mRNA for CD63 antigen was found in atherosclerotic lesions of Watanabe heritable hyperlipidemic rabbits, suggesting a potential role of CD63 in progression of atherosclerosis. CD63 is also a mast cell marker.
- 10
- 15

CD82 was originally identified as the target of several mAbs inhibitory to

- 20 syncytium formation induced by human T-cell leukemia virus type I (HTLV-I), the etiological agent of adult T-cell leukemia. Therefore, this gene could be a target for the development of a drug for this leukemia. CD81 is the target of an antiproliferative antibody. A diverse group of human cell lines, including hematolymphoid, neuroectodermal, and mesenchymal cells, express the CD81 protein. Many of the
- 25 lymphoid cell lines, in particular those derived from large cell lymphomas, were susceptible to the antiproliferative effects of the antibody. CD81 may therefore play an important role in the regulation of lymphoma cell growth. CD9, CD20, CD37, CD63, CD81 and CD82 have been implicated in the regulation of cell growth, adhesion, and signal transduction of B, T lymphocytes and some other non-lymphoid
- 30 cells. They associate with CD2, CD21, CD4, CD8, MHC Class II molecules, integrins, function as co-receptor for T, B and other lymphoid cells. Some

TM4SF are leukocyte antigens, highly expressed in activated leukocytes, lymphocytes, are highly specific surface marker for lymphoblastic leukemia, lymphoma, melanoma, and neuroblastoma. CD9 has been shown to be involved in cell motility and tumor metastasis. These antigen could be a valuable immunogen or

5 target to implement active and passive immunotherapy in patients with cancer. Others have been shown to be involved in inhibition of prostate cancer metastasis. This gene has close homology to C33 antigen (CD82), which is a member of the transmembrane 4 superfamily (TMSF) and activation antigen of T- cells. C33 Ag (CD82) was originally identified as the target of several mAbs inhibitory to syncytium formation

10 induced by human T-cell leukemia virus type I (HTLV-I), the etiological agent of adult T-cell leukemia. Therefore, this gene could be very important target for developing drug for leukemia. Other members of this family are Sm23, CO-029, R2, TAPA-1, CD9, CD37, CD53, and CD63. CD63 is a 53-Kd lysosomal membrane glycoprotein that has been identified as a platelet activation molecule.

15 There is strong evidence indicating that CD63 and Pltgp40, a platelet membrane glycoprotein are the same molecule and that CD63/Pltgp40 is identical to the well-characterized, stage-specific melanoma-associated antigen ME491. These antigen could be valuable immunogens or target to implement active and passive immunotherapy in patients with cancer.

20 This gene is expressed primarily in fetal tissue (kidney, heart, liver, spleen, brain), macrophages, dendritic cells, retina and to a lesser extent in various other tissues, mostly of lymphoid origin or epithelial cell types. In addition This gene is expressed in cancerous tissues (e.g. breast).

Therefore, polynucleotides and polypeptides of the invention are useful as
25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic diseases and/or disorders and cancers in a variety of organs and cell types. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential
30 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at

significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., developmental, proliferating, immune, hematopoietic, integumentary, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid, spinal fluid, or amniotic fluid) or another tissue or cell sample taken

- 5 from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 133 as residues: Tyr-123 to Tyr-131, Cys-134 to Ser-

- 10 145, Tyr-234 to Tyr-244. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution fetal cells and tissues and homology to tumor antigens indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, treatment and diagnosis of lymphoid and epithelial disorders and

- 15 neoplasms. Additionally, tissue distribution in immune cells and other tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting hematopoiesis, including cancers. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere 20 herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

- 25 Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, 30 such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity

disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits

- 5 hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types.

Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of 10 cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern 15 formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of 20 potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types 25 of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new 30 insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue

markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
10 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1918 of SEQ ID NO:41, b is an integer of 15 to 1932, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

- The translation product of this gene shares limited sequence homology with VEGF which is thought to be important in regulation of endothelial cell growth. Therefore, it is likely that the protein encoded by this gene would share some similar
20 biological functions.

- When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other immune and hematopoietic cells and tissue cell types, through the Jak-STAT signal transduction pathway. The gamma activating
25 sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and
30 differentiation of cells.

This gene is expressed in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, nervous system disease and/or disorders. Similarly, polypeptides and

- 5 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily
- 10 fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic

- 15 epitopes shown in SEQ ID NO: 134 as residues: Thr-25 to Pro-46. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions.

- 20 Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia,
- 25 trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene
- 30 product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to

- 5 identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

- 10 related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
15 formula of a-b, where a is any integer between 1 to 1150 of SEQ ID NO:42, b is an integer of 15 to 1164, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 33

- 20 The translation product of this gene shares sequence homology with human p150 which is thought to be important in signal transduction in neuronal cells. Therefore, it is likely that the protein encoded by this polynucleotide would share some similar biological functions with p150.

This gene is expressed primarily in whole embryo and cerebellum.

- 25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological and growth defects/disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological
30 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS, expression of this gene

at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

- 5 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of central nervous system, neurodevelopmental, cognitive, and memory disorders. The tissue distribution

- 10 also indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not
15 limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive
20 disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation,

- 25 neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show
30 utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked

by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders"

- 5 and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent
10 of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the
15 polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in
20 modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their
25 interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
30 related to SEQ ID NO:43 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

- excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1091 of SEQ ID NO:43, b is an integer of 15 to 1105, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:43, and where b is greater than or equal to a + 14.
- 5

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

- This gene is expressed primarily in PMA stimulated HL-60 cells and to a lesser extent in 6 week embryo.
- 10
- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders affecting cell differentiation, particularly hematopoietic
- 15
- disorders and/or defects. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the metabolic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types
- 20
- (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.
- 25
- Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 136 as residues: Pro-61 to Asp-68. Polynucleotides encoding said polypeptides are also provided.
- The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study of cellular differentiation and for
- 30
- the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia. The tissue distribution also

indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the

- 5 expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses). Moreover, the protein may
- 10 represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Additionally, the expression within embryonic tissue and other
- 15 cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere
- 20 herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain

- 25 neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating,
- 30 detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue

differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new

- 5 insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or
- 10 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:44 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

- 15 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1248 of SEQ ID NO:44, b is an integer of 15 to 1262, where both a and b correspond to the positions of nucleotide
- 20 residues shown in SEQ ID NO:44, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 35

This gene is expressed primarily in colon.

- Therefore, polynucleotides and polypeptides of the invention are useful as
- 25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders and/or defects of the digestive tract including but not limited to cancers of the gastrointestinal tract. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential
 - 30 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, expression of this gene at

significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., gastrointestinal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

- 5 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the digestive system particularly disorders involving the colon. Further, expression of this

- 10 gene product in colon tissue indicates involvement in digestion, processing, and elimination of food, as well as a potential role for this gene as a diagnostic marker or causative agent in the development of colon cancer, and cancer in general. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the colon and/or other gastrointestinal tissue
- 15 including, but not limited to, stomach, small intestine, large intestine, and rectum.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:45 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

- 20 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 503 of SEQ ID NO:45, b is an integer of 15 to 517, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:45, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 36

This gene is expressed primarily in blood cells.

- Therefore, polynucleotides and polypeptides of the invention are useful as
30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, immune and hematopoietic diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic

- 5 system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or
- 10 bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 138 as residues: Pro-19 to Cys-29, Thr-35 to Glu-44, Val-72 to Lys-78. Polynucleotides encoding said polypeptides are also provided.

- The tissue distribution indicates that polynucleotides and polypeptides
- 15 corresponding to this gene are useful for treatment and diagnosis and/or treatment of disorders of the immune and hematopoietic system. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or
- 20 activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

- Many polynucleotide sequences, such as EST sequences, are publicly
- 25 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:46 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
- 30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 844 of SEQ ID NO:46, b is an

integer of 15 to 858, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 37

5 This gene is expressed in multiple tissue systems such as brain, immune cells, prostate, uterus, testes, placenta, and fetal heart as well as in cancerous tissues such as ovarian tumors. .

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a 10 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the immune, reproductive, urogenital, and central nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, 15 particularly of the central nervous system and immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, reproductive, urogenital, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative 20 to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 139 as residues: Tyr-33 to Lys-38. Polynucleotides encoding said polypeptides are also provided.

25 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the immune, urogenital, reproductive, and central nervous systems. The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis 30 of diseases of the central nervous system, as well as cancers of tissues where expression of this gene has been observed, such as in ovarian tumors. The tissue

distribution in central nervous system tissues also indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome,

- 5 schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo. . Protein, as well as,
- 10 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and
- 15 disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

- Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in

proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution in uterus indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating female infertility. The protein product is likely involved in preparation of the endometrium of implantation and could be administered either topically or orally. Alternatively, this gene could be transfected in gene-replacement treatments into the cells of the endometrium and the protein products could be produced. Similarly, these treatments could be performed during artificial insemination for the purpose of increasing the likelihood of implantation and development of a healthy embryo. In both cases this gene or its gene product could be administered at later stages of pregnancy to promote healthy development of the endometrium. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution in testes indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that is expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product is expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Protein, as well as, antibodies directed

against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
- 5 related to SEQ ID NO:47 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
- 10 formula of a-b, where a is any integer between 1 to 6093 of SEQ ID NO:47, b is an integer of 15 to 6107, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:47, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 38

- 15 This gene is expressed in a wide range of tissue systems such as brain, immune cells, fetal liver, kidney, testes, breast, and pancreas as well as cancerous tissue such as ovarian tumors.
- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
- 20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the central nervous system, immune system, urogenital, and reproductive system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above
- 25 tissues or cells, particularly of the immune and central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, CNS, urogenital, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a
- 30 disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 140 as residues: Met-1 to Ser-7, Asp-32 to Pro-43, Ser-96 to Arg-102. Polynucleotides encoding said polypeptides are also provided.

- The tissue distribution indicates that polynucleotides and polypeptides
- 5 corresponding to this gene are useful for treatment and diagnosis of disorders of the immune, reproductive, urogenital and central nervous systems. The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and/or diagnosis of diseases of the central nervous system, as well as cancers of tissues where
- 10 expression of this gene has been observed, such as in ovarian tumors. The tissue distribution in central nervous system tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome,
- 15 schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders.
- 20 Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation
- 25 of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).
- Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such
- 30 as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities,

such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's

- 5 Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the
- 10 protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within
- 15 embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below
- 20 and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

- Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain
- 25 neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating,
 - 30 detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue

differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases.

The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The

- 5 protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show
10 utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
15 related to SEQ ID NO:48 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
20 formula of $a-b$, where a is any integer between 1 to 689 of SEQ ID NO:48, b is an integer of 15 to 703, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where b is greater than or equal to $a + 14$.

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

- 25 This gene is expressed primarily in macrophages and fetal cells and to a lesser extent in cancerous ovarian tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

- 30 not limited to, immune diseases, disorders of developing tissues, and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in

providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal and immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, cancerous

- 5 and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides

- 10 corresponding to this gene are useful for treatment and diagnosis of developmental abnormalities and disorders of the immune systems. The tissue distribution cancerous ovaries indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and intervention of these tumors. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker
- 15 and/or immunotherapy target for the above listed tissues. Expression of this gene product in macrophage cells strongly indicates a role for this protein in immune function and immune surveillance. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). This gene
- 20 product may have clinical utility in the treatment of immune dysfunction; in the correction of autoimmunity; in immune modulation; and in the control of inflammation.

The tissue distribution indicates polynucleotides and polypeptides

- corresponding to this gene are useful for the diagnosis and treatment of a variety of
25 immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene
30 product is involved in the regulation of cytokine production, antigen presentation, or

other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for

- 5 immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity
- 10 disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful
- 15 in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types.

Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention

- 20 of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

- 25 Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have
- 30 applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the

polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of

- 5 degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue
- 10 markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution also indicates that polynucleotides and polypeptides corresponding to this gene are useful
- 15 for the treatment, diagnosis, and/or prevention of various skin disorders such as melanomas.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:49 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 625 of SEQ ID NO:49, b is an integer of 15 to 639, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 40

- This gene is expressed primarily in neutrophils, bone marrow, brain, and fetal
- 30 cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic disorders, Limbic system dysfunction/defects and

- 5 disorders of the immune system and developing systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, Limbic system and developing systems, expression of this gene at significantly higher or lower levels is
10 routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 15 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 142 as residues: Ala-84 to Gln-93. Polynucleotides encoding said polypeptides are also provided.

- The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the
20 immune, Limbic system, CNS and developing systems. Expression of this gene product in bone marrow, eosinophils, and neutrophils strongly indicates a role for this protein in hematopoiesis and immune surveillance. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune
25 responses). This gene product may have clinical utility in the treatment of immune dysfunction; in the correction of autoimmunity; in immune modulation; and in the control of inflammation. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution indicates polynucleotides and polypeptides
30 corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune

"Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene

- 5 product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for

- 10 immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity
15 disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury.

- 20 Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to
25 its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Additionally, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment,
30 and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the

"Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell

- 5 death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may
- 10 also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of
- 15 degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue
- 20 markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly

- 25 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
- 30 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 853 of SEQ ID NO:50, b is an

integer of 15 to 867, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:50, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

5 This gene is expressed primarily in ovary and to a lesser extent in fetal tissue, colon, and immune cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

10 not limited to, ovarian cancer, gastrointestinal and immune system disorders.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the female reproductive system, expression of this gene at significantly higher or lower

15 levels is routinely detected in certain tissues or cell types (e.g., reproductive, gastrointestinal, immune, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an

20 individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 143 as residues: Ile-23 to Ala-29. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of ovarian cancer and related metastases. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating female infertility. The tissue distribution in colon tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders involving the gastrointestinal tract. This may include diseases associated with digestion and food absorption, as well as hematopoietic disorders involving the

Peyer's patches of the small intestine, or other hematopoietic cells and tissues within the body. Similarly, expression of this gene product in colon tissue indicates again involvement in digestion, processing, and elimination of food, as well as a potential role for this gene as a diagnostic marker or causative agent in the development of

- 5 colon cancer, and cancer in general. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment,
10 and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

- 15 Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have
20 applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue
25 differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the
30 protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their

interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly

- 5 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
- 10 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1555 of SEQ ID NO:51, b is an integer of 15 to 1569, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 42

The translation product of this gene shares sequence homology with retrovirus-related reverse transcriptase pseudogene. In addition, this gene shares homology with human interferon-beta (Genseq accession number T35524; all references available through this accession are hereby incorporated herein by reference), therefore, it is likely that this gene and the protein encoded by this gene shares some similar biological functions with this protein.

This gene is expressed primarily in frontal cortex.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
- 25 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system,
 - 30 expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily

fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 The tissue distribution in frontal cortex and homology to retrovirus-related reverse transcriptase pseudogene and human interferon-beta indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neurodegenerative diseases of the brain, particularly of the frontal cortex. The tissue distribution indicates polynucleotides and polypeptides
- 10 corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of
- 15 Alzheimer's Disease, Parkinson's Disease, multiple sclerosis, cystic fibrosis, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder,
- 20 learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

- Potentially, this gene product is involved in synapse formation,
- 25 neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show
- 30 utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

- 5 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1182 of SEQ ID NO:52, b is an integer of 15 to 1196, where both a and b correspond to the positions of nucleotide
- 10 residues shown in SEQ ID NO:52, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 43

This gene is expressed primarily in immune cells, brain, fetal tissue, and cancerous tissues (such as testes, stomach, lung, pancreas, ovaries) and to a lesser

- 15 extent in other numerous tissues including, but not limited to, testes and kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative diseases. Similarly, polypeptides and antibodies

- 20 directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and immune cells expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids
- 25 (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 30 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 145 as residues: Lys-23 to Lys-35, Met-46 to Tyr-52. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neurodegenerative disorders of the frontal cortex, as well as, cancer or a number of tissues including but not limited to testes, stomach, lung, pancreas, and ovaries. The

- 5 tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but
- 10 are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation,

- 20 neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.
- 25 The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere
- 30 herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell

lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

- Since the gene is expressed in cells of lymphoid origin, the natural gene
- 5 product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and
- 10 tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits
- 15 hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their
- 20 interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility
- 25 in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.
- 30 Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent

of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may

- 5 also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of
- 10 degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue
- 15 markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly
- 20 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
 - 25 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 931 of SEQ ID NO:53, b is an integer of 15 to 945, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where b is greater than or equal to a + 14.

30 **FEATURES OF PROTEIN ENCODED BY GENE NO: 44**

This gene is expressed primarily in epithelioid sarcoma and to a lesser extent in pancreatic carcinoma, aorta endothelial cells induced with TNF-alpha, and amniotic cells induced with TNF. This gene is also expressed, to a lesser extent, in cancerous lung and ovary tissue and fetal tissue.

- 5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, epithelioid sarcoma and related cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological
- 10 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., amniotic, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another
- 15 tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 146 as residues: Tyr-39 to Arg-51. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of certain cancers, including epithelioid sarcoma and pancreatic carcinoma. The tissue distribution in tumors of lung, ovary, and pancreas origins indicates that polynucleotides and

25 polypeptides corresponding to this gene are useful for the diagnosis and intervention of these tumors, in addition to other tumors where expression has been indicated. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Moreover, the expression within embryonic tissue and other cellular sources marked

30 by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of

- developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.
- Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including

blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

- Since the gene is expressed in cells of lymphoid origin, the natural gene
- 5 product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and
- 10 tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits
- 15 hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their
- 20 interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
- 25 related to SEQ ID NO:54 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
- 30 formula of a-b, where a is any integer between 1 to 474 of SEQ ID NO:54, b is an

integer of 15 to 488, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:54, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 45

5 In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

PPVPPWISLPLTGSPRPGFVPVSPFCFSPTNGHQVLLLLLTSAVAAGPWPQ

10 VHAGQWGWMCLPPGLPSVQARSGLGGLPGGPQWVPGGARGY (SEQ ID NO: 234). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in fetal and infant tissue, particularly infant brain and fetal liver/spleen libraries, and to a lesser extent in breast, ovary tumor, pharynx carcinoma, endometrial stromal cells, thymus, islet cell tumors, and adult

15 cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and other proliferative disorders. Similarly, polypeptides and

20 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and breast, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, developmental, hematopoietic, and cancerous and 25 wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 The tissue distribution in developing cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis

and treatment of cancer and other proliferative disorders. The expression within cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and

- 5 other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell

- 10 death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may
15 also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of
20 degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue
25 markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly

- 30 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:55 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

- 5 formula of a-b, where a is any integer between 1 to 2846 of SEQ ID NO:55, b is an integer of 15 to 2860, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:55, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

- 10 In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

IQQWGDSVLGRRCRDLLLQLYLQRPEL RVPVPEVLLHSEGAASSSVCKLDGLI
15 HRFITLLADTSR ALENRGADASMACRKLAVAHPLLLRLHLP MIAALLHGR
THLN FQEFRQQNHLSCFLHV LGLELLQPHVFRSEHQ GALWDCLLSFIRLLL
YRKSSRHLAAFINKFVQFIHKYITYNAPAAISFLQKHADPLHDLSFDNSDLVM
LKSLLAGLSLPSRDDRTDRGLDEEGEEESSAGSLPLVS VSLFTPLTAAEMAPY
MKRLSRGQTVEDLLEVLS DIDE MSRRP EILSFFSTNLQRLMSSAEECCR NLA
20 FSLALRSMQNSPSIAAAFLPTF MYCLGSQDFEVVQTALRN LPEY ALLCQEHA
AVLLHRAFLVGMY GQM DPSAQISEALRILHMEAVM (SEQ ID NO: 235).

Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in breast cancer, and to a lesser extent in a variety of other cancers, including uterine cancer, synovial sarcoma, and pharynx carcinoma.

- 25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, breast cancer; proliferative diseases and/or disorders. Similarly,
30 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For

a number of disorders of the above tissues or cells, particularly of the breast, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, breast, proliferative, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, breast milk, urine,

- 5 synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic

- 10 epitopes shown in SEQ ID NO: 148 as residues: Glu-35 to His-41, Ser-62 to Ala-67, Pro-145 to Leu-155, Glu-157 to Ser-163, Arg-190 to Val-197, Asp-208 to Pro-215, Ser-247 to Pro-252. Polynucleotides encoding said polypeptides are also provided.

- The tissue distribution in breast cancer tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of cancer. Elevated expression of this gene product in cancers, such as breast cancer, suggest that it is involved in the abnormal proliferation of cells, dedifferentiation, angiogenesis, and other processes that accompany the development of cancer. Thus, therapeutics targeted against this gene product is useful therapeutic products in and of themselves. Alternately, expression of this gene product at elevated levels in breast tissue is reflective of expression within breast lymph nodes, and may suggest a hematopoietic role for this protein. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

- 25 Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have 30 applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the

polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of

- 5 degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue
- 10 markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly

- 15 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
- 20 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1545 of SEQ ID NO:56, b is an integer of 15 to 1559, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:56, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 47

The translation product of this gene shares limited sequence homology with cytochrome-c oxidase. An alternative embodiment is the polypeptide comprising the following amino acid sequence:

MLLKHLQRMVSPQVKASALKVVTLTANDKTSVSFSSLPGQGVIYNVIVWD

- 30 PFLNTSAAAYIPAHTYACSFEAGEGSCASLGRVSSKVFFTLFALLGFFICFFGHR
FWKTELFFIGFIIMGFFFYILITRLTPIKYDVNLILTAVTGSVGGMFLVAWWWR

FGILSICMLCVGLVLGFLISSVTFFTPLGNLKIFHDDGVFWVTFSCIAILIPVVF
MGCLRILNILTCAVGIVGSYSVVLAIDSYWSTSLSYITLNVLKRALNKDFHRAFTN
VPFQTNDFIILA VWGMLAVSGITLQIRRERGRPFPPHPYKLWKQERERRVTNI
LDPSYHIPPLRERLYGRLTQIKGLFQKEQPAGERTPLLL (SEQ ID NO: 236).

5 In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

WARLRGPGAHARTSPQPWRGPSPAQAAMGFLQLLVVXLSEHRVAGAAE
10 VFGNSS EGLIEFSVGKF RYF
ELNRPFPPEEAILHD ISSNVTF LIFQIHSQYQNTTVSFSPRRRSPTM (SEQ ID NO:
237). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in keratinocytes, brain, and spinal cord and to a lesser extent in hematopoietic cells and tissues.

15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative disorders; hematopoietic disorders; integumentary disorders; immune dysfunction; learning disabilities. Similarly, polypeptides and
20 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and nervous systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., integumentary, neural, developmental,
25 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 The tissue distribution in brain and spinal cord cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the

diagnosis and treatment of a variety of neurological and hematopoietic disorders. For example, elevated levels of expression of this gene product in brain and spinal cord indicates that it is involved in neurodegenerative disorders. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in

5 Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms,

10 hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

15 Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Alternately, expression of this gene product in hematopoietic cells indicates that it is involved in the proliferation, differentiation, survival, and activation of all hematopoietic lineages, including stem and progenitor cells. Expression of this gene

20 product in keratinocytes indicates that it is involved in normal skin function, and could be involved in skin disorders, dermatitis, and fibrosis. The protein is useful in detecting, treating, and/or preventing congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's Disease, basal cell carcinoma, squamous cell carcinoma,

25 malignant melanoma, Paget's Disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and

30 xanthelasma. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox,

molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athletes foot, and ringworm). Moreover, the protein product of this gene may also be useful for the treatment or diagnosis of various connective tissue disorders (i.e., arthritis, trauma, tendonitis, chondromalacia and inflammation, etc.), autoimmune

- 5 disorders (i.e., rheumatoid arthritis, lupus, scleroderma, dermatomyositis, etc.), dwarfism, spinal deformation, joint abnormalities, and chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate
10 cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly

- 15 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
20 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2050 of SEQ ID NO:57, b is an integer of 15 to 2064, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:57, and where b is greater than or equal to a + 14.

25 **FEATURES OF PROTEIN ENCODED BY GENE NO: 48**

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

- 30 PRVRPASPPVRSPARWGSMAGSPLLWGPRAGGVGLLVLLLLGLFRPPPALCA
RPVKEPRGLSAASPPLARLALLAASGGQCPEVRRRGRCRPGAGAGASAGAER

QERARAEAQRLRISRRASWRSCCASGAPPATLIRLWAWTTPTRLQRSSLALC
SAPALTLPP (SEQ ID NO: 238). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in human pituitary and to a lesser extent in
5 pineal gland, and other areas of the brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pituitary dysfunction; abnormal growth; neurological defects;

- 10 insufficient milk secretion; abnormal smooth muscle contraction. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and nervous systems, expression of this gene at significantly higher or lower levels is
- 15 routinely detected in certain tissues or cell types (e.g., endocrine, developmental, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, breast milk, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or
- 20 bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 150 as residues: Pro-36 to Gly-42, Pro-64 to Ala-76, Gly-83 to Ala-90, Ser-100 to Cys-108, Thr-126 to Ser-135. Polynucleotides encoding said polypeptides are also provided.

- 25 The tissue distribution primarily in pituitary cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of disorders. Elevated expression of this gene product in the pituitary indicates that it is possibly a hormone-like substance that either controls pituitary development itself, or various processes controlled by the
- 30 pituitary. These include growth, milk secretion, smooth muscle contraction, diuresis, blood pressure, and homeostasis. Thus, this gene product may have numerous clinical

applications. Expression of this gene product in other regions of the brain also indicates that it is involved in normal neurological function, and is useful in the treatment of a variety of neurological disorders. Representative uses are described in the "Biological Activity", "Hyperproliferative Disorders", and "Binding Activity"

- 5 sections below, in Example 11, 17, 18, 19, 20 and 27, and elsewhere herein. Briefly, the protein can be used for the detection, treatment, and/or prevention of Addison's Disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-,hypoparathyroidism) ,
10 hypothalamus, and testes. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility

- 15 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
20 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1036 of SEQ ID NO:58, b is an integer of 15 to 1050, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:58, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the

- 30 following amino acid sequence:

PRVRLATPNIWDLMLFAFISLLVMLPTWWIVSSWLVWGVILFVYLVIRALRL

WRTAKLQVTLKKYSVHLEDMATNSRAFTNLVRKALRLIQETEVISRGFTLVS
AACPFNKAGQHPSQHLIGLRKAVYRTLQRNFQAARLATLYMLKNYPLNSES
DNVTNYICVVVPFKELGLGLSEEQISEEEAHNFTDGFSLPALKVLFQLWVAQSS
EFFRRLLALLLSTANSPPGPLLTPALLPHRILSDVTQGLPHAHSACLEELKRSYE
5 FYRYFETQHQSVPQCLSQTQQKSRELNNVHTAVRSLQLHLKALLNEVIILEDE
LEKLVCTKETQELVSEA YPILEQKLKLIQPHVQASNNCWEAISQVDKLLRRN
TDKKKGKPEIACENPHCTVSTFEAAYSTHCRQRSNPRGAGIRSLCR (SEQ ID
NO: 239). Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have a transmembrane
10 domain at about amino acid position 7 - 23 of the amino acid sequence referenced in
Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 24 to
390 of this protein has also been determined. Based upon these characteristics, it is
believed that the protein product of this gene shares structural features to type Ib
membrane proteins.

15 The gene encoding the disclosed cDNA is believed to reside on chromosome
12. Accordingly, polynucleotides related to this invention are useful as a marker in
linkage analysis for chromosome 12.

This gene is expressed primarily in prostate and placenta and to a lesser extent
in pancreatic tumors and hematopoietic cells.

20 Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, prostate cancer; pancreatic cancer; prostate dysfunction; hematopoietic
disorders; reproductive diseases and/or disorders, and pancreatitis. Similarly,
25 polypeptides and antibodies directed to these polypeptides are useful in providing
immunological probes for differential identification of the tissue(s) or cell type(s). For
a number of disorders of the above tissues or cells, particularly of the endocrine and
immune systems, expression of this gene at significantly higher or lower levels is
routinely detected in certain tissues or cell types (e.g., reproductive, prostate,
30 pancease, placental, vascular, and cancerous and wounded tissues) or bodily fluids
(e.g., lymph, serum, seminal fluid, plasma, urine, synovial fluid and spinal fluid) or

another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 151 as residues: Pro-85 to Ser-94, Pro-127 to Thr-136, Glu-154 to Glu-160, Phe-240 to Ser-250, Leu-255 to Leu-265, Leu-341 to Lys-351, Thr-372 to Gly-384. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in prostate and placental cells and tissues indicates that 10 polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of reproductive disorders. Elevated expression of this gene product in the prostate indicates that it is involved in normal prostate function, and is a diagnostic marker for prostate cancer. Alternately, expression of this gene product in placenta indicates that it may play a role in normal vascular 15 function, and is involved in such processes as angiogenesis and endothelial cell chemotaxis. Thus, this gene product is useful in the treatment of myocardial infarction, cancer, ischemia, and diabetic retinopathy. Expression of this gene product in placenta may also be indicative of fetal health and development.

Similarly, expression of this gene product in hematopoietic cells indicates that 20 it is involved in the proliferation, differentiation, survival, or activation of all hematopoietic cell lineages. Finally, expression of this gene product in pancreatic cancers indicates that it may play a role in cancer in general, or in pancreatic function. The secreted protein can also be used to determine biological activity, to raise 25 antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, and as nutritional supplements. It may also have a very wide range of biological activities. Representative uses are described in the "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Briefly, the protein may possess the 30 following activities: cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases

- and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:59 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2519 of SEQ ID NO:59, b is an integer of 15 to 2533, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:59, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

When tested against Jurkat and K562 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) and ISRE (interferon-sensitive responsive element) promoter elements, respectively. Thus, it is likely that this gene activates myeloid, leukemia, and to a lesser extent, other immune or hematopoietic cells and tissue cell-types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.

Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. ISRE is also a promoter element found upstream in many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a

- 5 large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

In another embodiment, polypeptides comprising the amino acid sequence of
10 the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

AAPHPLLRLPLCLWCPLWPAWPLRGRPRSAWKRWPPLPVGPAKLGCSMTTR
QPTAVSWPCWLMSSSLSTACLAWTLTGSLAREATRRARSLSPWNCSARQV
15 PPSPPHSGLGRRGWAHCHLT CLLVTQLFRVGRIHPILSPLV (SEQ ID NO:
240). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in brain and placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, vascular diseases; aberrant angiogenesis; neurological disorders; learning disorders; placental insufficiency; and fetal distress. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of
25 disorders of the above tissues or cells, particularly of the vascular and neurological systems (CNS/PNS), expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, reproductive, vascular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an
30 individual having such a disorder, relative to the standard gene expression level, i.e.,

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 152 as residues: Met-1 to Thr-7, Glu-36 to Ser-43,

- 5 Pro-46 to Gly-63. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in brain and placental cells and tissues, combined with the detected GAS and ISRE biological activities, indicates that the protein products of this gene are useful for the diagnosis and/or treatment of a variety of neural, reproductive, and vascular diseases and/or disorders. neurodegenerative disease

- 10 states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis,

- 15 encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and

- 20 perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Expression of this gene product in placenta indicates that it may play a role

- 25 in blood vessel development or function, as the placenta is a highly vascularized organ. Thus, this gene product is involved in such processes as angiogenesis, endothelial cell chemotaxis, and vascular cord formation. Thus, it is useful in the treatment of such conditions as myocardial infarction; ischemia; and cancer.

- Alternately, expression of this gene product in the brain indicates that it may play a 30 role in the survival, proliferation, or function of neurons, and thus is useful in the diagnosis and treatment of such neurological disorders as ALS, schizophrenia, and

Alzheimer's Disease. It may likewise be involved in learning disorders as well. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement.

- 5 Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:60 and may have been publicly available prior to conception of

- 10 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 885 of SEQ ID NO:60, b is an
15 integer of 15 to 899, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:60, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 51

- In another embodiment, polypeptides comprising the amino acid sequence of
20 the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

LQLASQSAGIKGMSHCARPTFLLLASCFWAAAIPNRNVILSVSFRPLHMQ

FTLSILVFILRILILLRSFL (SEQ ID NO: 241). Polynucleotides encoding these

- 25 polypeptides are also provided.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 40 - 56 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 57 to 60 of this protein has also been determined. Based upon these characteristics, it is
30 believed that the protein product of this gene shares structural features to type Ia membrane proteins.

This gene is expressed primarily in spleen derived from patients with chronic lymphocytic leukemia.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

- 5 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, chronic lymphocytic leukemia; hematopoietic disorders; impaired immune function; cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above
- 10 tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard
- 15 gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in spleen cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of a variety of hematopoietic disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

- 20 The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Elevated expression of this protein in the spleens of patients with CLL indicates that it is a useful marker for thi's
- 25 Disease. Alternately, it is associated with the development and/or progression of the disease, and is a useful target for therapeutic intervention. Additionally, this gene
- 30

product may play more general roles in hematopoiesis, and may serve to control cellular decisions regarding proliferation, survival, activation, and/or differentiation of all hematopoietic cell lineages. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate

- 5 ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly

- 10 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:61 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or
15 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1065 of SEQ ID NO:61, b is an integer of 15 to 1079, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:61, and where b is greater than or equal to a + 14.

20 **FEATURES OF PROTEIN ENCODED BY GENE NO: 52**

The translation product of this gene shares sequence homology with a putative tyrosine protein kinase from the Chilo iridescent virus. See, for example, Genbank accession no. gi|2738451 (AF003534). Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities
25 with tyrosine kinase and signaling proteins. Such activities are known in the art, some of which are described elsewhere herein.

This gene is expressed in a variety of tissues, including microvascular endothelial cells, dendritic cells, and fetal tissues, as well as several tumors.

- Therefore, polynucleotides and polypeptides of the invention are useful as
30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

not limited to, vascular, immune, and developmental diseases and/or disorders, particularly cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above

- 5 tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., vascular, immune, developmental, proliferative, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having
- 10 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 154 as residues: Ala-21 to Lys-31, Arg-41 to Cys-56, Thr-92 to Cys-102, Arg-132 to Val-137, Lys-152 to Ile-159, Pro-199 to Ser-205, Arg-15 210 to Asp-219, Ser-225 to Lys-230, Tyr-236 to Ala-241, Lys-243 to Leu-249, Thr-375 to Asp-381. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution and homology to a tyrosine kinase indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of cancer. Representative uses are described in the "Immune Activity" 20 and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved 25 in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such 30 as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities,

such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's

- 5 Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the
- 10 protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or embolism. For example, this gene product may represent a soluble factor produced by smooth muscle that regulates the innervation of organs or regulates the survival of neighboring neurons.
- 15 Likewise, it is involved in controlling the digestive process, and such actions as peristalsis. Similarly, it is involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.
- 20
- 25

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:62 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1914 of SEQ ID NO:62, b is an

integer of 15 to 1928, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:62, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 53

5 The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 2 - 18 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

This gene is expressed primarily in neutrophils.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic diseases and/or disorders, particularly cancer and immune suppression. Similarly, polypeptides and antibodies directed to 15 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids 20 (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic 25 epitopes shown in SEQ ID NO: 155 as residues: Gly-63 to Ser-72. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful as a marker for neutrophil monitoring in cancer and/or immune suppressed patients and/or during chemotherapy 30 or radiation therapy. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and

elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes

- 5 suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such

- 10 as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic
15 lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages,
20 and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or
25 immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:63 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically
30 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 767 of SEQ ID NO:63, b is an integer of 15 to 781, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:63, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 54

This gene is expressed primarily in IL-1 and LPS induced neutrophils, and to a lesser extent, in fetal brain.

- Therefore, polynucleotides and polypeptides of the invention are useful as
- 10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hematopoietic, and neural diseases and/or disorders, particularly cancer and immune suppression. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for
- 15 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, neural, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, amniotic fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.
- 20

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 156 as residues: Ile-28 to Trp-37, Ser-68 to Lys-81.

- 25 Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful as a marker in neutrophils to monitor patients who are immune suppressed or cancer patients during chemotherapy or radiation therapy. Representative uses are described in the "Immune Activity" and

30 "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in

regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune

5 responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory

10 bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's

15 Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively,

20 polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in

25 addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

30 related to SEQ ID NO:64 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1180 of SEQ ID NO:64, b is an

- 5 integer of 15 to 1194, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:64, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 55

This gene is expressed primarily in prostate.

- 10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, urogenital diseases and/or disorders, particularly prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in
- 15 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the urogenital system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., urogenital, prostate, renal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,
- 20 synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 157 as residues: Arg-30 to Gln-36. Polynucleotides encoding said polypeptides are also provided.

- The tissue distribution in prostate cancer cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, treatment and diagnosis of prostate cancer and other urogenital disorders. Moreover, the expression 30 within cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis,

treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

- Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:65 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1663 of SEQ ID NO:65, b is an integer of 15 to 1677, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:65, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 56

A preferred polypeptide of the invention comprises the following amino acid sequence:

MVLVLRHPLCARERAFAREPGRGLLRTGQHDGAPAVTAVPGPLGAVAAAEG

- 10 RRSAWGAGGSSPPRKVLWGDMRGRAGVDVLGPALSSEAAGAEARGWGM
PGMGVGVGASETRGALFLGREGVHGPCPMGLGPWPWGPW (SEQ ID NO:
242). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in rejected kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as

- 15 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders affecting the kidney. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of
20 disorders of the above tissues or cells, particularly of the urinary tract, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., urogenital, renal, kidney, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative
25 to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 158 as residues: Ala-30 to Gly-36, Asp-45 to Trp-50, Lys-65 to Cys-71, Pro-80 to Cys-87. Polynucleotides encoding said polypeptides are
30 also provided.

The tissue distribution in kidney indicates the protein product of this gene could be used in the treatment and/or detection of kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, 5 renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. The protein is useful for modulating the immune response to aberrant proteins, as may exist in proliferating cells and tissues. Such modulation of the immune response would also show utility in inhibiting the rejection of transplanted 10 tissues, particularly of the renal system. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the 15 above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:66 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically 20 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1223 of SEQ ID NO:66, b is an integer of 15 to 1237, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:66, and where b is greater than or equal to a + 14. 25

FEATURES OF PROTEIN ENCODED BY GENE NO: 57

The translation product of this gene shares sequence homology with both human and mouse Fibulin-2 which is an extracellular matrix protein found in heart 30 tissue (See Genbank Accession Nos. emb|CAA57876.1 and emb|CAA53040.1, respectively; all references available through these accessions are hereby incorporated

herein by reference; for example, J. Cell Biol. 123 (5), 1269-1277 (1993)). Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

MGPAVKMWTNAWKGLDDCHYNQLCENTPGGHRCSCPGRGYRMQGPSLPCL
DVNECLQLPKACAYQCHNLQGSYRCLCPPGQTLLRDGKACTSLERNGQNVT

- 5 TVSHRGPLLWLRPWASIPGTSYHAWVSLRPGPMALSSVGRAWCPPGFIRQN
GVCTDLDECVRNLCQHACRNTEGSYQCLCPAGYRLLPSGKNCQDINECEEEE
SIECGPGQMCNFNTRGSYQCVDTCPATYRQGPSPGTCFRRCSQDCGTGGPSTL
QYRLLPLPLGVRAHHDVARLTAFSEVGVPANRTESMLEPDPRSPFALRPLRA
GLGAVYTRRALTRAGLYRLTVRAAAPRHQSVFVLLIAVSPYPY (SEQ ID NO:
10 243). Polynucleotides encoding these polypeptides are also provided.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence:

MRVLVVTIAPIYWALARESGEALNGHSLTGGKFRQSHTWSLLQGAHDDPV
ARGLDPDGLLLLDVVVNVVPGRAWLTQIFKCRTLKKHYVQTRAWPAVRG

- 15 LHTALLPGRPPLVPTLQPQHPVQRGPAPAGAAPAGLSYQLGL (SEQ ID
NO: 244). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the

- 20 following amino acid sequence:

HASGAFLVVRGEHQGSMTGVINGRKFGVATLNTSVMQEAHSGVSSIHSS
IRHVPANVGPLMRVLVVTIAPIYWALARESGEALNGHSLTGGKFRQESHVEF
ATGELLTMTQWPGVWIPMASCSSTWWMSMALSPDSLADADLQVQDFEEHYV
QTGPQLFVGSTQRFFQGGLPSFLRCNHISIQYNAARGPQPQLVQHLRASAIISS

- 25 AFDPEAEALRFQLATALQAEENEVGCPEGFELDSQGAFCVDVDECAWD AHL
CREGQRCVNLLGSYRCLPDCPGFDRVADGAGCEDVDECLEGLDDCHYNQLC
ENTPGGHRCSCPGRGYRMQGPSLPCLDVNECLQLPKACAYQCHNLQGSYRCL
CPPGQTLLRDGKACTSLERNGQNVTIVSHRGPLLWLRPWASIPGTSYHAWV
SLRPGPMALSSVGRAWCPPGFIRQNGVCTDLDECVRNLCQHACRNTEGSY
30 QCLCPAGYRLLPSGKNCQDINECEEEE SIECGPGQMCNFNTRGSYQCVDTCPAT
YRQGPSPGTCFRRCSQDCGTGGPSTLQYRLLPLPLGVRAHHDVARLTAFSEV

GVPANRTELSMLEPDPRSPFALRPLRAGLGAVYTRRALTRAGLYRLTVRAAA
PRHQSVFVLLIAVSPYPY (SEQ ID NO: 245). Polynucleotides encoding these polypeptides are also provided.

- When tested against U937 and Jurkat cell lines, supernatants removed from
- 5 cells containing this gene repeatedly activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid, T-cells, and to a lesser extent, other immune and hematopoietic cells and tissue cell types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway
- 10 is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

- This gene is expressed primarily in kidney.
- 15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and/or disorders affecting the kidney and renal system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in
- 20 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the urinary tract, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., renal, urogenital, kidney, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,
- 25 synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 159 as residues: Lys-32 to Ser-37, His-89 to Gly-94, Asn-124 to Gln-130, Ala-163 to Val-168, Cys-196 to Arg-201, Gln-244 to Gln-264,

His-288 to Tyr-294, Leu-314 to Gln-319, Ala-392 to Ser-399, Pro-412 to Asp-419, Ala-452 to Pro-460, Arg-466 to Thr-473. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in rejected kidney, the homology to the conserved

- 5 Fibulin-2 protein, in addition to the detected GAS biological activity, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting kidneys, particularly proliferative disorders. Representative uses are described here and elsewhere herein. The protein product of this gene could be used in the treatment and/or detection of kidney diseases
10 including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Furthermore, the protein may also be used
15 to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 20 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:67 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
25 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1920 of SEQ ID NO:67, b is an integer of 15 to 1934, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:67, and where b is greater than or equal to a + 14.

Preferred polypeptides of the invention comprise the following amino acid sequence:

MGEKFLLLAMKENHPECFCKILKILHCMDPGEWLPQTEHCVHLTPKEFLIWT
MDIASNERSEIQSVALRLASKVISHHMQTCVENRELIAELKQWVQLVILSCE

5 DHLPTESRALAVVEVLSTTPLFLTNPHPILELQDTLALWKCVLTLQSEEQAV
RDAATETVTTAMSQEENTCQSTEFACQVDASIALALALA VLCDLLQQWDQL
APGLPILLGWLLGESDDL VACVESMHQVEEDYLFEKAEVNFWAETLIFVKYL
CKHLFCLLSKSGWRPPSPEMLCHLQRMVSEQCHLLSQFFRELPPAAEFVKTV
EFTRLRIQEERTLA CLRLLAFLEGKEGEDTLVLSVWDSYAESRQLTLPRTEAA

10 C (SEQ ID NO: 246). Polynucleotides encoding such polypeptides are also provided.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MGEPNRHPSM

FLLLLVLERLYASPMDGTSSALSMGPVFPFIMRCGHSPVYHSREMAARALVP
FVMIDHIPNTIRTLLSTL

15 PSCTDQCFCRAKPHSGHFSRFHLLQAYSDSKTRNEFRLPARAD (SEQ ID NO:
247). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the

20 following amino acid sequence:

MTGREFFSRFPELYPFLLKQLETVANTVDSDMGEPNRHPSMFLLLLVLERLY
ASPMDGTSSALSMGPVFPFIMRCGHSPVYHSREMAARALVPFVMIDHIPNTIR
TLLSTLPSCTDQCFCRNHIHGTLQVFHLLQAYSDSKHGTNSDFQHELTIDITV
CTKAKLWLAKRQNPNCLVTRAVYIDILFLLCCLNRSAKDNQPVLESLGFWEET

25 VRGIISGS ELITGFPWAFKVPGLPQYLQSLTRLAIAAVWAAA AKSGERETNVPI
SFSQLLES AFPEVRSLTLEALLEKFLAAASGLGEKGVPPLL CNMGEKFLLLAM
KENHPECFC KILKILHCM DPGEWLPQTEHCVHLTPKEFLIWTMDIASNERSEIQS
VALRLASKVISHHMQTCVENRELIAELKQWVQLVILS CEDHLPTESRALAVV
EVLTSTTPLFLTNPHPILELQDTLALWKCVLTLQSEEQA VRDAATETVTTAM
30 SQENTCQSTEFACQVDASIALALALA VLC DLLQQWDQLAPGLPILLGWLLG
ESDDL VACVESMHQVEEDYLFEKAEVNFWAETLIFVKYLCKHLFCLLSKSG

WRPPSPEMLCHLQRMVSEQCHLLSQFFRELPPAAEFVKTVEFTRLRIQEERTL
ACLRLLAFLEGKEGEDTLVLSVWDSYAESRQLTPRTEAAC (SEQ ID NO:
248). Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have two transmembrane domains at about amino acid position 144 - 160, and 462 - 478 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type IIIa membrane proteins. Included in this invention as a preferred domain is the formate and nitrite transporters domain, which was identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). A number of bacterial and archaeabacterial proteins involved in transporting formate or nitrite have been shown [1] to be related:

- focA and focB, from Escherichia coli, transporters involved in the bidirectional transport of formate.
- fdhC, from Methanobacterium formicum and thermoformicum, a probable formate transporter.
- nirC, from Escherichia coli and Salmonella typhimurium, a probable nitrite transporter.
- Bacillus subtilis hypothetical protein yrhG.
- Bacillus subtilis hypothetical protein ywcJ (ipa-48R).

These transporters are proteins of about 280 residues and seem to contain six transmembrane regions. As signature patterns, we selected two conserved regions. The first one is located in what seems to be a cytoplasmic loop between the second and third transmembrane domains; the second is part of the fourth transmembrane region. The 70 Kd yeast hypothetical protein YHL008c is highly similar, in its N-terminal section, to the prokaryotic members of this family. The consensus pattern is as follows: [LIVMA]-[LIVMY]-x-G-[GSTA]-[DES]-L-[FI]-[TN]-[GS].

Preferred polypeptides of the invention comprise the following amino acid sequence: IISGSELITG (SEQ ID NO: 249). Polynucleotides encoding these polypeptides are also provided. Further preferred are polypeptides comprising the formate and nitrite transporter domain of the sequence referenced in Table for this gene, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of this referenced sequence. The additional contiguous amino acid residues is 25 N-terminal or C-terminal to the formate and nitrite transporter domain. Alternatively, the additional contiguous amino acid residues is both N-terminal and C-terminal to 30

the formate and nitrite transporter domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domain is characteristic of a signature specific to formate and nitrite transporter proteins. Based on the sequence similarity, the translation product of this

5 gene is expected to share at least some biological activities with formate and nitrite transporter proteins. Such activities are known in the art, some of which are described elsewhere herein. It is believed that this gene maps to chromosome 2. Accordingly, polynucleotides derived from this gene are useful in linkage analysis as markers for chromosome 2.

10 This gene is expressed primarily in cells of the immune system, primarily T-cells and to a lesser extent in spleen, liver, thymus, tonsils, and testis.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

15 not limited to, immune and hematopoietic diseases and/or disorders, particularly disorders affecting hematopoiesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of hematopoietic cells, expression of this gene at

20 significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from

25 an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 160 as residues: Gly-2 to Pro-8, Ser-82 to His-92, Tyr-107 to Asp-117, Arg-162 to Pro-169, Ser-224 to Thr-229, Leu-310 to His-315, Ser-333 to Glu-338, Glu-381 to Ser-388, Gln-428 to Ala-433, Met-446 to Thr-455,

30 Ser-548 to Ser-554, Gly-613 to Asp-618, Ser-627 to Gln-633. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in immune cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting hematopoiesis, including cancers.

Representative uses are described in the "Immune Activity" and "infectious disease"

- 5 sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a

10 usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory

- 15 bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's
- 20 Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the
- 25 protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 30 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

related to SEQ ID NO:68 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or

- 5 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3286 of SEQ ID NO:68, b is an integer of 15 to 3300, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:68, and where b is greater than or equal to a + 14.

10 **FEATURES OF PROTEIN ENCODED BY GENE NO: 59**

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

- 15 VDGIDKLDIEFLQQFLETHSRGPRLHSPGHASQEATPGANMSSGTELLWPGAA
LLVLLGVAASLCVRCRPGAKRSEKIYQQRSLRDQQSFTGSRTYSLVGQAW
PGPLADMAPTRKDQLQFYPSLEDPASSRYQNFSKGSRGSEEAYIDPIAMEY
YNWGRFSKPPEDDDANSYENVLICKQKTETGAQQEGIGGLCRGDLSLAL
KTGPTSGLCPSASPEEDEGI (SEQ ID NO: 250). Polynucleotides encoding these
20 polypeptides are also provided.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 10 - 26 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ib membrane proteins.

- 25 The gene encoding the disclosed cDNA is believed to reside on chromosome 7. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 7.

- 30 This gene is expressed primarily in bone marrow, CD34 positive cells, and immune cells, including, neutrophils, T-cells, B-cells, macrophages, monocytes, and dendritic cells and to a lesser extent in brain and tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders affecting the immune and hematopoietic systems, particularly

- 5 hematopoiesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the the immune system and hematopoeitic system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or
- 10 cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 15 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 161 as residues: Ser-29 to Thr-57, Pro-74 to Lys-79, Pro-85 to Glu-107, Tyr-118 to Tyr-136, Gln-144 to Gln-152, Ala-182 to Glu-188. Polynucleotides encoding said polypeptides are also provided.

- 20 The tissue distribution in immune and hematopoietic cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of disorders affecting the immune system and hematopoiesis. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in
- 25 regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).
- 30 Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for

immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and

- 5 tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits
- 10 hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia,
- 15 thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or
- 20 receptors, to identify agents that modulate their interactions, and as nutritional supplements. It may also have a very wide range of biological activities.

Representative uses are described in the "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15, 16, 18, 19, and 20, and elsewhere herein. Briefly, the protein may possess the following activities: cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for 25 treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors);

- 30

hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the

- 5 corresponding nucleic acid in gene therapy procedures.

Based upon the the proteins immune cell specific message distribution, it may be involved

in many aspects of the immune response, especially its initial stages, inflammation, allograft rejection, infectious disease response etc. The expression of this clone is

- 10 frequently found in the hematopoietic cell cDNA libraries. Thus, this factor could be involved in the control of hematopoietic cell proliferation, differentiation, and function. Based on this one can postulate its use in the management of anemias, leukemias, neutropenia, thrombocytopenia, autoimmune diseases, blood tissue engraftment, and poikilothromerythromatosis. Furthermore, the
15 protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 20 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:69 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is
25 cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1783 of SEQ ID NO:69, b is an integer of 15 to 1797, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:69, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 60

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

- 5 VLWREASALVLSNRLSSGLLHDLLLQPAIHISRLFPRRSRGLSEGEGRSSVSLQRS
RVLSAMKHVLNLYLLGVVLTLLSIFVRVMESLEGLLESPSPGTSWTTRSQLAN
TEPTKGLPDHPSRSM (SEQ ID NO: 251). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in immune cells including activated T cells, 10 macrophages, jurkat cells, bone marrow cells, and osteoblasts and to a lesser extent in kidney cortex, brain, placenta and lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are 15 not limited to, immune and hematopoietic diseases and/or disorders, particularly inflammation and diseases related to inflammatory activity. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression 20 of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in 25 healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 162 as residues: Pro-34 to Met-63. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in immune cells and tissues indicates that 30 polynucleotides and polypeptides corresponding to this gene are useful for treating or diagnosing disease related to the normal or abnormal activation of T cells.

Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell

- 5 lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for
10 immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity
15 disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful
20 in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,
25 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:70 and may have been publicly available prior to conception of
30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1359 of SEQ ID NO:70, b is an integer of 15 to 1373, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:70, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by 10 the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

YTFHTQIFLDFPMIFLTFLPLAFLFLHSGFYHYISFSCLFSLSLALFFFLDVATFR
RPGQLFCERSVLFDMFHFGFVSLFLHEWIQAKHFWAGLF
IVLPSDVFFSVHHLEAPDGSPNIAKLSLIILLR (SEQ ID NO: 252).

15 Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have two transmembrane domains at about amino acid position 2 - 18 and 22 - 38 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type IIIa membrane 20 proteins.

This gene is expressed in many tissues including brain, liver, prostate, testes, cartilage, gall bladder. Expression is also seen in a number of tumors including colon carcinoma, pancreas tumor, osteoclastoma, ovarian cancer, B cell lymphoma and acute lymphocytic leukemias.

25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, tumors of various organs including the pancreas, colon, and bone. Similarly, polypeptides and antibodies directed to these polypeptides are useful in 30 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the

major organs, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, hepatic, metabolic, reproductive, testicular, skeletal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or

- 5 another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in tumors and proliferative tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating or

- 10 diagnosing tumors of several major organs including the pancreas and large intestine. This protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below

- 15 and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain

- 20 neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may also act as a morphogen to control cell and tissue type specification. Therefore, the

- 25 polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in

- 30 proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the

protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or

- 5 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:71 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

- 10 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1565 of SEQ ID NO:71, b is an integer of 15 to 1579, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:71, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 62

This gene is expressed primarily in dendritic cells and fetal liver/spleen and to a lesser extent in many tissues including tonsils, fetal lung, stromal cell lines, bone

- 20 marrow cell lines, placenta and tumors including hepatocellular carcinoma, pancreas tumor and osteosarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

- 25 not limited to, diseases and/or disorders of the immune and hematopoietic system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels is
- 30 routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,

synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 The tissue distribution in dendritic cells and fetal liver/spleen indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosing and treating disorders of the immune system particularly related to the control and generation of precursor cells. polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic
- 10 related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow
- 15 transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

- The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of
- 20 stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies
- 25 directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:72 and may have been publicly available prior to conception of
- 30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1014 of SEQ ID NO:72, b is an integer of 15 to 1028, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:72, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 63

This gene is expressed primarily in adrenal gland tumor and endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as

- 10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endocrine and vascular diseases and/or disorders, particularly diseases associated with the vascular endothelium. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for
- 15 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., endocrine, vascular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or
- 20 cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 25 The tissue distribution in endothelial cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosing and treating disorders that involve the vascular system including diseases such as atherosclerosis, neoangiogenesis associated with tumor growth and conditions associated with inflammation. Moreover, the protein is useful in the detection, treatment, and/or prevention of a variety of vascular disorders and conditions, which include, but are not limited to microvascular disease, vascular leak syndrome, aneurysm, stroke, embolism, thrombosis, coronary artery disease, arteriosclerosis, and/or atherosclerosis. Alternatively, the protein is useful in the treatment, detection, and/or
- 30

prevention of metabolic disorders, particularly lethargy and depression. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,

- 5 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:73 and may have been publicly available prior to conception of

- 10 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3660 of SEQ ID NO:73, b is an
15 integer of 15 to 3674, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:73, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

- The translation product of this gene is related to bovine PAM precursor. See
20 Genbank record gi|163482 incorporated herein by reference. Moreover, see following patent publications are also incorporated herein by reference: J04311386 and WO8902460. Many bioactive peptides terminate with an amino acid alpha-amide at their COOH terminus. The enzyme responsible for this essential posttranslational modification is known as peptidyl-glycine alpha-amidating monooxygenase or PAM.
25 An NH2-terminal signal sequence and short propeptide precede the NH2 terminus of purified PAM. The sequences of several PAM cyanogen bromide peptides were localized in the NH2-terminal half of the predicted protein. The forms of PAM purified from bovine neurointermediate pituitary is generated by endoproteolytic cleavage at a subset of the 10 pairs of basic amino acids in the precursor. High levels
30 of PAM mRNA have been found in bovine pituitary and cerebral cortex. In

corticotropic tumor cells, levels of PAM mRNA and pro-ACTH/endorphin mRNA are known to be regulated in parallel by glucocorticoids and CRF.

This gene is expressed primarily in endometrial tumors, dendritic cells, a multiple sclerosis library, kidney, hematopoietic cells, melanocytes, osteoblasts, the spleen, colon, ovary, stromal cells, fetal and adult brain, heart, and in tissues undergoing wound repair.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endometriosis, endometrial cancer, multiple sclerosis, hematopoietic diseases, bone disease, and wound healing. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly the hematopoietic system and female reproduction, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., reproductive, immune, hematopoietic, integumentary, skeletal, gastrointestinal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- The tissue distribution in dendritic and hematopoietic cells and tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful as a therapeutic or diagnostic agent in Diseases of hematopoietic origin as well as the female reproductive track due to the gene's primary pattern of expression. Polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses

include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency

- 5 etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The protein may also have a very wide range of biological activities. Representative uses are described in the "Chemotaxis" and "Binding Activity" sections below, in Examples 11, 12, 13, 14, 15,
- 10 16, 18, 19, and 20, and elsewhere herein. Briefly, the protein may possess the following activities: cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's Disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.
- 15
- 20
- 25 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:74 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
- 30

formula of a-b, where a is any integer between 1 to 2783 of SEQ ID NO:74, b is an integer of 15 to 2797, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:74, and where b is greater than or equal to a + 14.

5 **FEATURES OF PROTEIN ENCODED BY GENE NO: 65**

The translation product of this gene shares sequence similarity with several G-protein coupled receptors (See Genbank Accession No. gb|AAC77910.1| (AF061443); all references available through this accession are hereby incorporated herein by reference; for example, Mol. Endocrinol. 12, 1830-1845 (1998)). G-protein coupled receptors are well known in the art and affect a variety of functions. In particular, the translation product of this gene shares similarity with Follicle Stimulating Hormone Receptor.

Preferred polypeptides encoded by this gene comprise the following amino acid sequence:

15 GTRFPTGETPSLGFTVTLVLLNSLAFLLMAVIYTKLYCNLEKEDLSENSQSSMI
KHVAWLIFTNCIFFCPVAFFSFAPLITAISISPEIMKSVTLIFFP (SEQ ID NO: 253). Polynucleotides encoding such polypeptides are also provided.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence: MIKHVAWLIFTNCIFFCP

20 VAFFSFAPLITAISISPEIMKSVTLIFFPCLLA (SEQ ID NO: 254). Polynucleotides encoding these polypeptides are also provided.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the

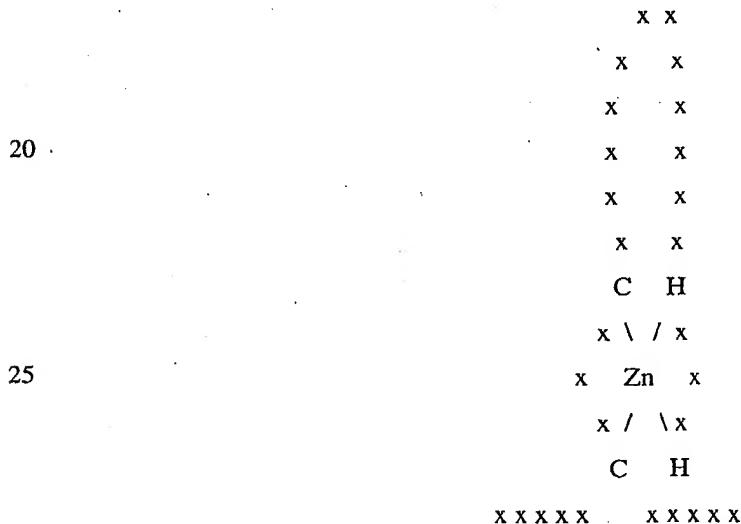
25 following amino acid sequence:

GTRFPTGETPSLGFTVTLVLLNSLAFLLMAVIYTKLYCNLEKEDLSENSQSSMI
KHVAWLIFTNCIFFCPVAFFSFAPLITAISISPEIMKSVTLIFFPLPACLNPVLYVF
FNPKFEDWKLLKRRVTKKSGSVSVSISSQGGCLEQDFYYDCGMYSHLQGN
LTVCDCCESFLLTkpVsCKHLIKSHSCPALAVASCQRPEGYWSDCGTQSAHS

30 DYADEEEDSFVSDSSDQVQACGRACFYQSRGFPLVRYAYNLPRVKD (SEQ ID NO: 255). Polynucleotides encoding these polypeptides are also provided.

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 43 - 59 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 60 to 207 of this protein has also been determined. Based upon these characteristics, it is

- 5 believed that the protein product of this gene shares structural features to type Ia
membrane proteins. Included in this invention as preferred domains are Zinc finger,
C2H2 type domains, which were identified using the ProSite analysis tool (Swiss
Institute of Bioinformatics). 'Zinc finger' domains [1-5] are nucleic acid-binding
protein structures first identified in the *Xenopus* transcription factor TFIIIA. These
10 domains have since been found in numerous nucleic acid-binding proteins. A zinc
finger domain is composed of 25 to 30 amino-acid residues. There are two cysteine or
histidine residues at both extremities of the domain, which are involved in the
tetrahedral coordination of a zinc atom. It has been proposed that such a domain
interacts with about five nucleotides. A schematic representation of a zinc finger
15 domain is shown below:



- 30 Many classes of zinc fingers are characterized according to the number and positions of the histidine and cysteine residues involved in the zinc atom

coordination. In the first class to be characterized, called C2H2, the first pair of zinc coordinating residues are cysteines, while the second pair are histidines. A number of experimental reports have demonstrated the zinc- dependent DNA or RNA binding property of some members of this class. Some of the proteins known to include

- 5 C2H2-type zinc fingers are listed below. We have indicated, between brackets, the number of zinc finger regions found in each of these proteins; a '+' symbol indicates that only partial sequence data is available and that additional finger domains is present. In addition to the conserved zinc ligand residues it has been shown that a number of other positions are also important for the structural integrity of the C2H2
- 10 zinc fingers. The best conserved position is found four residues after the second cysteine; it is generally an aromatic or aliphatic residue. The concensus pattern is as follows: C-x(2,4)-C-x(3)-[LIVMFYWC]-x(8)-H-x(3,5)-H.

- Preferred polypeptides of the invention comprise the following amino acid sequence: CDCCESFLLTKPVSKHLIKSH (SEQ ID NO: 256). Polynucleotides encoding these polypeptides are also provided. Further preferred are polypeptides comprising the Zinc finger, C2H2 type domain of the sequence referenced in Table for this gene, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of this referenced sequence. The additional contiguous amino acid residues is N-terminal or C- terminal to the Zinc finger, C2H2 type domain.
- 20 Alternatively, the additional contiguous amino acid residues is both N-terminal and C-terminal to the Zinc finger, C2H2 type domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domain is characteristic of a signature specific to zinc finger proteins. Based on the sequence similarity, the translation product of this gene is expected to
 - 25 share at least some biological activities with G-coupled proteins, their receptors, and zinc finger proteins. Such activities are known in the art, some of which are described elsewhere herein.

- This gene is expressed primarily in adult and fetal liver, human placenta, colon carcinoma cell lines and fibroblasts and to a lesser extent in the fetal and adult
- 30 brain, the developing nervous system, lung, pancreas, salivary gland, breast tissue, and dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the liver, developmental abnormalities, neurologic diseases,

- 5 lung cancer, pancreatic cancer, and colon cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neurological and hepatic origin, as well as the proliferation and/or differentiation of numerous types of tissues.
- 10 expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., hepatic, immune, hematopoietic, neural, gastrointestinal, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative
- 15 to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 167 as residues: Pro-62 to Asp-67, Arg-74 to Gly-80, Gln-146 to Glu-168. Polynucleotides encoding said polypeptides are also provided.

- 20 The tissue distribution in fetal liver indicates that polynucleotides and polypeptides corresponding to this gene are useful for a diagnostic marker or therapeutic in a wide variety of disease states. Polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or
- 25 leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of
- 30 neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the

- 5 differentiation and/or proliferation of various cell types. Alternatively, the protein expression in placental and brain tissue indicates the protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or embolism. For example, this gene product may represent a soluble
- 10 factor produced by smooth muscle that regulates the innervation of organs or regulates the survival of neighboring neurons. Likewise, it is involved in controlling the digestive process, and such actions as peristalsis. Similarly, it is involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. The protein is useful in the treatment, detection, and/or prevention of
- 15 bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's Disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic
- 20 depression, delirium, severe mental retardation and dyskinesias, such as Huntington's Disease or Gilles de la Tourette's syndrome. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed
- 25 against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:75 and may have been publicly available prior to conception of
- 30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2689 of SEQ ID NO:75, b is an integer of 15 to 2703, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:75, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 66

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by 10 the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

ALENSGSPGLQDSARAHFNXSLRSFSFLRNQMYIFELSLYLEGTSFVVVLLFLL
ISVSLDSPPTTKGWDSVLHIWVPLIVQ (SEQ ID NO: 257). Polynucleotides encoding these polypeptides are also provided.

15 This gene is expressed primarily in placenta and in hematopoietic cells, especially those of T-cell and monocyte origin and to a lesser extent in the brain, endothelial cells, and the lungs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a 20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic, vascular, and developmental diseases and/or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the 25 immune system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., vascular, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene 30 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 168 as residues: Ser-30 to Trp-37. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in hematopoietic cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for therapeutic and/or diagnostic intervention in hematopoietic and developmental disorders. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia.

The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the protein is useful in the detection, treatment, and/or prevention of vascular conditions, which include, but are not limited to, microvascular disease, vascular leak syndrome, aneurysm, stroke, atherosclerosis, arteriosclerosis, or embolism. For example, this gene product may represent a soluble factor produced by smooth muscle that regulates the innervation of organs or regulates the survival of neighboring neurons. Likewise, it is involved in controlling the digestive process, and such actions as peristalsis. Similarly, it is involved in controlling the vasculature in areas where smooth muscle surrounds the endothelium of blood vessels. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:76 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general

- 5 formula of a-b, where a is any integer between 1 to 728 of SEQ ID NO:76, b is an integer of 15 to 742, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:76, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

- 10 This gene is expressed primarily in the prostate and to a lesser extent in human B-cell lymphomas.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

- 15 not limited to, prostate cancer and diseases of hematopoietic origin, particularly of B-cells. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate and immune systems, expression of this gene at significantly higher or lower

- 20 levels is routinely detected in certain tissues or cell types (e.g., prostate, reproductive, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, seminal fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 169 as residues: Asp-33 to Lys-42. Polynucleotides encoding said polypeptides are also provided.

- 25 The tissue distribution in prostate tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful as a therapeutic or diagnostic marker for prostate cancer and disorders involving hematopoietic cells, especially

those of B-cell origin. Moreover, the expression within cellular sources marked by proliferating cells indicates this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, including cancer, and other proliferative

- 5 conditions. Representative uses are described in the "Hyperproliferative Disorders" and "Regeneration" sections below and elsewhere herein. Briefly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation.

Dysregulation of apoptosis can result in inappropriate suppression of cell
10 death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal muscular atrophy (SMA). Because of potential roles in proliferation and differentiation, this gene product may have applications in the adult for tissue regeneration and the treatment of cancers. It may
15 also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of
20 degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. The protein is useful in modulating the immune response to aberrant proteins and polypeptides, as may exist
25 in rapidly proliferating cells and tissues. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the
30 above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:77 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

- 5 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1811 of SEQ ID NO:77, b is an integer of 15 to 1825, where both a and b correspond to the positions of nucleotide
- 10 residues shown in SEQ ID NO:77, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 68

- When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of
- 15 the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.
 - 20

- In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the
- 25 following amino acid sequence:

GHEVICGSCRSWIYFSIRCRRMRPWWSLLEACATCAQTGPTRSTSCTQEVS
HSSSTAYPAPMRRRCCL PSPRSCT (SEQ ID NO: 258). Polynucleotides encoding these polypeptides are also provided.

- The gene encoding the disclosed cDNA is believed to reside on chromosome
- 30 17. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17.

This gene is expressed primarily in the brain and the developing embryo and to a lesser extent in the heart, colon, adipose tissue, kidney, mammary tissue, activated T-cells and dendritic cells.

- Therefore, polynucleotides and polypeptides of the invention are useful as
- 5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological diseases, developmental conditions, colon cancer, and hematopoietic diseases, especially of T-cell origin . Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological
- 10 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., neural, developmental, cardiovascular, adipose, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g.,
- 15 lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 170 as residues: Thr-18 to Cys-26, Glu-29 to Thr-36, Ser-50 to Thr-55. Polynucleotides encoding said polypeptides are also provided.

- The tissue distribution in brain, combined with the detected GAS biological activity, indicates that polynucleotides and polypeptides corresponding to this gene are useful for therapeutic and/or diagnostic agents in neurological diseases,
- 25 developmental abnormalities, colon cancer, and hematopoietic diseases, especially those of T-cell origin. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease,
- 30 Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal

cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated

- 5 expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity,

- 10 to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:78 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or 20 more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1660 of SEQ ID NO:78, b is an integer of 15 to 1674, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:78, and where b is greater than or equal to a + 14.

25 **FEATURES OF PROTEIN ENCODED BY GENE NO: 69**

The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 2 - 18 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein

- 30 product of this gene shares structural features to type II membrane proteins.

In another embodiment, polypeptides comprising the amino acid sequence of the open reading frame upstream of the predicted signal peptide are contemplated by the present invention. Specifically, polypeptides of the invention comprise the following amino acid sequence:

- 5 KragvevGGLVMALAGSVFVLGGVLVLCVERNGEGEMGWPQHLPKSQPLS
PPVAVRRCFSFERSWIDLLVETSSSMVTCRQQVGTPNGMEGRGGGPKTTPIRL
QLSGACAVRPEIQWEV (SEQ ID NO: 259). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in activated monocytes, dendritic cells, and
10 in the tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and hematopoietic diseases and/or disorders, particularly
15 leukemia, lymphomas, tumors of hematopoietic origin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic system, expression of this gene at significantly higher or lower levels is routinely detected in
20 certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.
25 Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 171 as residues: Gln-30 to Leu-38, Asn-75 to Thr-86. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution in activated monocytes, dendritic cells, and tonsils indicates that polynucleotides and polypeptides corresponding to this gene are useful
30 as a therapeutic and/or diagnostic agent for leukemias, lymphomas, and other diseases associated with cells of hematopoietic origin. Representative uses are described in the

"Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene

- 5 product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for
10 immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity

- 15 disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful
20 in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as,
25 antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:79 and may have been publicly available prior to conception of
30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is

cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2177 of SEQ ID NO:79, b is an integer of 15 to 2191, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:79, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 70

- When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates myeloid cells, and to a lesser extent, other immune cells and tissue cell types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.
- Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

The gene encoding the disclosed cDNA is believed to reside on chromosome 12. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 12.

This gene is expressed primarily in the placenta, brain, and liver and to a lesser extent in most other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic, neurological, vascular, and developmental diseases and/or disorders, particularly cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and nervous systems, expression of this gene at significantly higher or lower levels is routinely detected in certain tissues or

cell types (e.g., hematopoietic, neurological, vascular, developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, bile, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression

- 5 level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in brain tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful therapeutic and/or diagnostic agent in a multitude of disease states, particularly those involving the immune and

- 10 neurologic systems. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in Example 11, 15, and 18, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating
15 diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated
20 expression of this gene product in regions of the brain indicates it plays a role in normal neural function.

- Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the protein is useful in the detection, treatment, and/or prevention
25 of a variety of vascular disorders and conditions, which include, but are not limited to microvascular disease, vascular leak syndrome, aneurysm, stroke, embolism, thrombosis, coronary artery disease, arteriosclerosis, and/or atherosclerosis. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents
30 that modulate their interactions, in addition to its use as a nutritional supplement.

Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are
- 5 related to SEQ ID NO:80 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general
- 10 formula of a-b, where a is any integer between 1 to 1321 of SEQ ID NO:80, b is an integer of 15 to 1335, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:80, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 71

- 15 The translation product of this gene shares sequence homology with the murine Fig1 (interleukin-four induced gene 1) which shares homology to the monoamine oxidases, particularly in domains responsible for FAD binding. Preferred polypeptides encoded by this gene comprise the following amino acid sequence:
QDWKAERSQDPFEKCMQDPDYEQLLKVTLADNRIGGRIFTYRDQXTGWIG
- 20 ELGAMRMPSSHRILHKLCQGLGLNLTKFTQYDKNTWTEVHEXKLRNYVVEK
VPEKLGYALRPQEKGHSPEDIYQMALNQALKDLKALGCRKAMKKFERHTLL
EYLLGEGNLRSRAVQLLGDVMSEDGFFYLSFAELRAXSCLSDRLQYSRIVG
GWDLLPRALLSSLSGLVLLNAPVVAMTQGPHDVHVQIETSPPARNLKVLKAD
VVLLTASGPRAVKRITFS (SEQ ID NO: 260), and/or
- 25 LPRHMQEALRRLHYVPATKVFLSFRRPFWREEHIEGGHSNTDRPSRMIFYPPP
REGALLLASYTWSAAAFAAGLSREEALRLALDDVAALHGPVVRQLWDGT
GVVKRWAEDQHSQGGFVVQXPALWQTEKDDWTVPYGRIFAGEHTAYPHG
WWVETAVKSALRAAIKINSRKGPASDTASPEGHASDMEGQGHVHGVAASSPSH
DLAKEEGS (SEQ ID NO: 261). Polynucleotides encoding such polypeptides are
- 30 also provided.

A preferred polypeptide fragment of the invention comprises the following amino acid sequence:

MAPLALHLLVLVPILLSLVASQDWKAERSQDPFEKCMQDPDYEQLLKVTL
EADNRIGGRIFTYRDQXTGWIGELGAMRMPSSHRLHKLCQGLGLNLTKFTQ

- 5 YDKNTWTEVHEXKLRLNYVVEKPEKLGYALRPQEKGHSPEDIYQMALNQA
LKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPAVQLLGDVMSEDGFFYL
SFAEALRAXSCLSDRLQYSRIVGGWDLLPRALLSSLSGLVLLNAPVVAMTQG
PHDVH
VQIETSPPARNLKVLKADVVLITASGPRAVKRITFSPRCPATCRRRCGGCTTCR
- 10 PPRCS (SEQ ID NO: 262). Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with monoamine oxidases, disintegrins, metalloproteinases, and apoptosis modulating proteins. Such activities are known in the art, some of which are described elsewhere herein.
Polynucleotides encoding these polypeptides are also provided.

- 15 The polypeptide of this gene has been determined to have a transmembrane domain at about amino acid position 235 - 251 of the amino acid sequence referenced in Table 1 for this gene. Moreover, a cytoplasmic tail encompassing amino acids 252 to 319 of this protein has also been determined. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type Ia
20 membrane proteins.

This gene is expressed primarily in hematopoietic cells, particularly in dendritic cells, and activated monocytes and to a lesser extent in T-cells, endothelial cells, and cells associated with ulcerative colitis.

- Therefore, polynucleotides and polypeptides of the invention are useful as
25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, leukemias, lymphomas, and diseases associated with antigen presenting cells, in addition to apoptosis dependant events. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological
30 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression

of this gene at significantly higher or lower levels is routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a

- 5 disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred polypeptides of the present invention comprise immunogenic epitopes shown in SEQ ID NO: 173 as residues: Gln-22 to Gln-44, Ala-90 to Gly-95, Lys-137 to Trp-146, Arg-171 to Asp-181, Glu-370 to Ser-380, Asp-447 to Gly-452,

- 10 Gln-463 to Trp-469, Asn-504 to Ala-510, Asp-512 to His-519, Ala-541 to Val-550, Asn-558 to His-566. Polynucleotides encoding said polypeptides are also provided.

The tissue distribution immune and hematopoietic cells and tissues, combined with the homology to the murine Fig 1 gene indicates that polynucleotides and polypeptides corresponding to this gene are useful as a therapeutic and/or diagnostic

- 15 agent for hematopoietic diseases, especially those associated with antigen presenting cells. Representative uses are described in the "Immune Activity" and "infectious disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell
20 lineages, including blood stem cells. This gene product is involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product is involved in immune functions. Therefore it is also useful as an agent for
25 immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous Disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity
30 disorders, such as autoimmune infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's

Disease, and scleroderma. Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages,

- 5 and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or
- 10 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:81 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

- 15 excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1853 of SEQ ID NO:81, b is an integer of 15 to 1867, where both a and b correspond to the positions of nucleotide
- 20 residues shown in SEQ ID NO:81, and where b is greater than or equal to a + 14.

		ATCC Deposit Nr and Date	NT SEQ ID NO: X	5' NT of Total Clone Seq.	3' NT of NT Seq. ID NO: X	5' NT of AA SEQ ID NO: Y	First AA of Signal Pep	Last AA of Signal Pep	First AA of Secreted Portion	Last AA of Secreted Portion
Gene No.	cDNA Clone ID	Vector	pSport	11 1113	1 1113	232 232	103 1	26 27	106	
1	HISCN02	209878	Lambda ZAP II	12 983	102 983	69 69	104 1	57 58	86	
2	HHGDM70	209878	Uni-ZAP XR	13 973	1 973	68 68	105 1	37 38	302	
3	HHPGO40	209878	Uni-ZAP XR	82 984	1 984	74 74	174 1	37 38	224	
3	HHPGO40	209878	Uni-ZAP XR	05/18/98	14 1458	1 1458	312 312	106 1	20 21	
4	HAMGG68	209878	pCMV Sport 3.0	15 2005	1 2005	251 251	107 1	22 23	189	
5	HAPOM49	209878	Uni-ZAP XR	05/18/98	83 2664	1 2664	448 448	175 1	1 2	
5	HAPOM49	209878	Uni-ZAP XR	05/18/98	16 943	1 933	62 62	108 1	38 39	
6	HBGBA69	209878	Uni-ZAP XR	05/18/98	17 1503	588 1480	290 290	109 1	26 27	
7	HBJFJ26	209878	Uni-ZAP XR	05/18/98					128	

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID NO: NT Seq. X	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of AA of Start Signal NO: Y	AA SEQ ID NO: Y	First AA of Codon Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
7	HBIFI26	209878 05/18/98	Uni-ZAP XR 84	1328	413	1305	591	176	1	20	21
8	HCEDH38	209878 05/18/98	Uni-ZAP XR 18	1512	1	1438	222	222	110	1	26
9	HDPOJ08	209878 05/18/98	pCMV Sport 3.0 19	1655	1	1655	159	159	111	1	18
10	HDPRX82	209878 05/18/98	pCMV Sport 3.0 20	2525	1	2525	128	128	112	1	32
11	HELGK31	209878 05/18/98	Uni-ZAP XR 21	1396	25	1334	209	209	113	1	29
11	HCNUA40	97898 02/26/97 209044 05/15/97	pBluescript 85	1342	949	1237	960	960	177	1	33
12	HFP CX64	209878 05/18/98	Uni-ZAP XR 22	1069	1	1069	181	181	114	1	28
12	HFP CX64	209878 05/18/98	Uni-ZAP XR 86	1154	84	1154	257	257	178	1	28
12	HCEBW71	209225 08/28/97	Uni-ZAP XR 87	1197	141	1197	257	257	179	1	28

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID	5' NT of Total NO: NT Seq.	3' NT of Clone Seq.	5' NT of Start Codon	AA SEQ ID NO: Pep Y	Last AA of Sig Pep	First AA of Sig Pep	Last AA of Secreted Portion
17	HBMCF37	209683 03/20/98	pBluescript	92	1487	79	1487	170	184	1
17	HFLQB16	209641 02/25/98	Uni-ZAP XR	93	1653	394	1637	413	185	1
17	HAGFY16	97923 03/07/97	Uni-ZAP XR	94	1830	87	1786	128	186	1
18	HRACJ35	209878 05/18/98	pCMV Sport 3.0	28	1863	8	1863	99	120	1
18	HAWAZ34	209141 07/09/97	pBluescript SK-	95	1134	472	1132	687	187	1
19	HTWDE26	209878 05/18/98	pSport1	29	1626	1	1626	68	121	1
19	HMHBGN40	97901 02/26/97	Uni-ZAP XR	96	1772	69	1772	129	188	1
		209047 05/15/97							30	31
20	HUSSIB13	209878 05/18/98	pSport1	30	605	1	605	172	122	1
21	HBAFA02	209877 05/18/98	pSport1	31	931	359	931	46	123	1

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID NO: X	NT Total Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of AA SEQ ID NO: Y	First AA of Signal Pep.	Last AA of Signal Pep.	First AA of AA SEQ ID NO: Y	Last AA of AA SEQ ID NO: Y	First AA of AA SEQ ID NO: Y	Last AA of AA SEQ ID NO: Y	First AA of AA SEQ ID NO: Y	Last AA of AA SEQ ID NO: Y	First AA of AA SEQ ID NO: Y	Last AA of AA SEQ ID NO: Y	First AA of AA SEQ ID NO: Y	Last AA of AA SEQ ID NO: Y	
22	H2CBT175	209877 05/18/98	pBluescript SK-	32	1407	1	1407	32	32	124	1	23	24	60						
23	HAGDQ42	209877 05/18/98	Uni-ZAP XR	33	1526	1	1526	126	126	125	1	18	19	248						
24	HBMCJ42	209877 05/18/98	pBluescript	34	1737	41	1580	244	244	126	1	44	45	248						
25	HDPBQ71	209877 05/18/98 3.0	pCMV Sport	35	2242	6	2242	24	24	127	1	33	34	612						
26	HCEIG71	209877 05/18/98	Uni-ZAP XR	36	2235	2	2235	28	28	128	1	25	26	447						
27	HEIHL48	209877 05/18/98	Uni-ZAP XR	37	2971	560	2557	629	629	129	1	16	17	291						
27	HSKCT36	209580 01/14/98	Uni-ZAP XR	98	1955	1	1955	31	31	190	1	18	19	184						
28	HISAQ04	209877 05/18/98	pSport1	38	1163	1	1163	61	61	130	1	21	22	78						
29	HIACB89	209877 05/18/98	pBluescript SK-	39	1932	28	1930	95	95	131	1	23	24	333						
30	HTECC05	209877 05/18/98	Uni-ZAP XR	40	881	1	881	27	27	132	1	15	16	164						

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID	Total NO: NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	First AA of Start Signal	Last AA of Stop Signal	First AA of Secreted Portion	Last AA of ORF
31	HBJLF01	209877 05/18/98	Uni-ZAP XR	41	1932	201	1931	217	133	1
32	HBXGP60	209877 05/18/98	ZAP Express	42	1164	1	1164	143	134	1
33	HCE5B20	209877 05/18/98	Uni-ZAP XR	43	1105	1	1105	237	135	1
34	HCMSSQ56	209877 05/18/98	Uni-ZAP XR	44	1262	1	1262	148	136	1
35	HCNAH57	209877 05/18/98	Lambda ZAP II	45	517	1	517	35	137	1
36	HCUEPP91	209877 05/18/98	ZAP Express	46	858	2	858	266	266	1
37	HDPCJ91	209877 05/18/98	pCMV Sport 3.0	47	6107	1	6107	131	131	1
38	HDPGK25	209877 05/18/98	pCMV Sport 3.0	48	703	1	703	345	345	1
39	HE2DY70	209877 05/18/98	Uni-ZAP XR	49	639	1	639	137	137	1
40	HE2NV57	209877 05/18/98	Uni-ZAP XR	50	867	1	867	99	99	1

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID NO: X	NT Total Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Signal Seq.	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of Secreted Portion
41	HETBR16	209877 05/18/98	Uni-ZAP XR	51	1569	1	1569	161	143	1	21	22
42	HFXDG13	209877 05/18/98	Lambda ZAP II	52	1196	1	1196	43	43	144	1	37
43	HFXKY27	209877 05/18/98	Lambda ZAP II	53	945	1	945	44	44	145	1	19
44	HHPEC09	209877 05/18/98	Uni-ZAP XR	54	488	1	488	71	71	146	1	19
45	HISAD54	209877 05/18/98	pSport1	55	2860	1	2860	172	172	147	1	19
46	HJBCY35	209877 05/18/98	pBluescript SK-	56	1559	93	1272	232	232	148	1	23
47	HKAEEA19	209877 05/18/98	pCMV Sport 2.0	57	2064	1	1909	83	83	149	1	21
48	HKGDL36	209877 05/18/98	pSport1	58	1050	1	1050	55	55	150	1	33
49	HLDBS43	209877 05/18/98	pCMV Sport 3.0	59	2533	1	2533	73	73	151	1	26
50	HLWADD92	209877 05/18/98	pCMV Sport 3.0	60	899	1	899	197	197	152	1	34

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID NO: X	5' NT of Total Clonc Seq.	3' NT of NT Seq.	5' NT of AA SEQ	First AA of ID	Last AA of Signal NO:	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
51	HLYBII5	209877 05/18/98	pSportI Vector	61 1079	1 1079	92 1079	92 153	1 153	22 92	23 92	23 60	
52	HMEJE05	209889 05/22/98	Lambda ZAP II	62 1928	1 1928	25 1928	25 154	1 154	30 121	31 155	31 392	
53	HNGIX55	209889 05/22/98	Uni-ZAP XR	63 781	1 781	121 1194	138 1194	138 121	19 156	20 156	20 74	
54	HNHEX30	209889 05/22/98	Uni-ZAP XR	64 1194	1 1194	138 1677	138 1677	138 236	19 157	20 157	20 81	
55	HPJBJ133	209889 05/22/98	Uni-ZAP XR	65 1677	1 1677	130 1237	130 1237	130 236	19 158	20 158	20 53	
56	HRABA80	209889 05/22/98	pCMV Sport 3.0	66 1237	1 1237	130 1934	130 1934	130 191	28 159	29 159	29 102	
57	HRACD80	209889 05/22/98	pCMV Sport 3.0	67 1934	1 1934	191 1958	191 1958	191 191	28 191	29 191	29 17	
57	HRACD80	209889 05/22/98	pCMV Sport 3.0	99 1958	1 1958	191 2729	191 2729	191 677	16 677	16 677	16 146	
58	HSLCX03	209889 05/22/98	Uni-ZAP XR	68 3300	984 3300	191 2444	984 2444	191 2444	16 392	16 392	16 23	
58	HSLCX03	209889 05/22/98	Uni-ZAP XR	100 2444	1 2444	192 2444	1 2444	192 2444	23 392	23 392	23 124	

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID NO: X	5' NT of Total Clone Seq.	3' NT of NT Seq.	5' NT of AA SEQ ID NO: Y	Last AA of Sig Pep	First AA of Sig Pep	Last AA of Sig Pep	First AA of Sig Pep	Last AA of Sig Pep	First AA of Sig Pep	Last AA of Sig Pep	First AA of Sig Pep	Last AA of Sig Pep
59	HT5GJ57	209889 05/22/98	Uni-ZAP XR	69	1797	92	1797	122	122	161	1	25	26	190	
60	HTACS42	209889 05/22/98	Uni-ZAP XR	70	1373	1	1373	213	213	162	1	29	30	63	
61	HTEKE40	209889 05/22/98	Uni-ZAP XR	71	1579	1	1579	173	173	163	1	47	48	117	
62	HTOBX69	209889 05/22/98	Uni-ZAP XR	72	1028	1	1028	28	28	164	1	20	21	42	
63	HUVB077	209889 05/22/98	Uni-ZAP XR	73	3674	1	3674	55	55	165	1	27	28	47	
64	H2CBG48	209889 05/22/98	pBluescript SK-	74	2797	1	2797	125	125	166	1	25	26	45	
65	H2CBU83	209889 05/22/98	pBluescript SK-	75	2703	1	2703	157	157	167	1	30	31	207	
65	H2CBU83	209889 05/22/98	pBluescript SK-	101	2709	1	2709	157	157	193	1	30	31	51	
66	HAPNY94	209889 05/22/98	Uni-ZAP XR	76	742	1	742	94	94	168	1	29	30	50	
67	HBJHZ58	209889 05/22/98	Uni-ZAP XR	77	1825	1	1825	102	102	169	1	29	30	42	

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	NT SEQ ID NO: X	5' NT of Total NT Seq.	3' NT of Clone Seq.	5' NT of Start Signal Seq.	5' NT of AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
68	HCE2B33	209889 05/22/98	Uni-ZAP XR Vector	78	1674	1	1668	67	67	170	1
69	HDPBQ02	209889 05/22/98	pCMV Sport 3.0	79	2191	291	2191	460	460	171	1
70	HFIY170	209889 05/22/98	pSportI	80	1335	1	1335	43	43	172	1
71	HDPOZ56	209889 05/22/98	pCMV Sport 3.0	81	1867	415	1867	103	103	173	1
71	HDPOZ56	209889 05/22/98	pCMV Sport 3.0	102	1722	1	1722	59	59	194	1

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related 5 DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits 10 contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." 15 and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is 20 identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal 25 peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

30 SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further

below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and

- 5 diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or

- 10 deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or
- 15 deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of

- 20 plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits.
- Moreover, the amino acid sequence of the protein encoded by a particular clone can
- 25 also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

- The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be
- 30 isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed

sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from 5 the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides 10 produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains 15 secretory or leader sequences, pro-sequences, sequences which aid in purification , such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced 20 version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

25

Signal Sequences

Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of 30 McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the

information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, *supra*.) However, the two
5 methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., *Protein Engineering* 10:1-6 (1997)), which predicts the cellular location of a
10 protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes
15 vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely
20 uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the
25 naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

30 **Polynucleotide and Polypeptide Variants**

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

- 5 By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the
- 10 polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire
- 15 sequence shown in Table 1, the ORF (open reading frame), or any fragment specified as described herein.

- As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 30 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject

- 5 sequences truncated at the 5' or 3' ends, relative to the the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then
- 10 subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the
- 15 purposes of manually adjusting the percent identity score.

- For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent
- 20 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the
 - 25 deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes
 - 30 of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95%

"identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other

- 5 words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those
- 10 terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be

- 15 determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the
- 20 query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size
- 25 Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

- 30 If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the the

query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by

- 5 results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query
- 10 sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C-termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are made for the purposes of the present invention.

- 30 The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter

the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred.

Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced

- 5 for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an

- 10 organism. (*Genes II*, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level.

Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA

- 15 technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., *J. Biol. Chem.* 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after

- 20 deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., *J. Biotechnology* 7:199-216 (1988).)

- 25 Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (*J. Biol. Chem.* 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1 α . They used random mutagenesis to generate over 3,500 individual IL-1 α mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every
- 30 possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See,

Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

- Furthermore, even if deleting one or more amino acids from the N-terminus or
- 5 C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular
- 10 polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions,

15 inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

20 The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these

25 positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function.

30 For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used.

(Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which

- 5 amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.
- 10
- 15 Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.
- 20
- 25 For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)
- 30

A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of the present invention having an amino acid sequence which contains at least one amino acid substitution, but not more than 50 amino acid substitutions, even more preferably, not more than 40 amino acid
5 substitutions, still more preferably, not more than 30 amino acid substitutions, and still even more preferably, not more than 20 amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a polypeptide to have an amino acid sequence which comprises the amino acid sequence of the present invention, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1
10 amino acid substitutions. In specific embodiments, the number of additions, substitutions, and/or deletions in the amino acid sequence of the present invention or fragments thereof (e.g., the mature form and/or other fragments described herein), is 1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, conservative amino acid substitutions are preferable.

15

Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about
20 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and
25 primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-
30 450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-

1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or

- 5 smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

20 Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

25 30 Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and

alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

- 5 Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to 10 an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

- 15 In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In 20 contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) 25 further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson 30 et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., *supra*; Wilson et al., *supra*; Chow, M. et al., *Proc. Natl. Acad. Sci. USA* 82:910-914; and Bittle, F. J. et al., *J. Gen. Virol.* 66:2347-2354 (1985).) A preferred immunogenic epitope includes

- 5 the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes
10 in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear

- 15 more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., *J. Nucl. Med.* 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

20

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the

- 25 polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

- 30 Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous

functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life *in vivo*. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., *Nature* 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., *J. Biochem.* 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See,

D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide.

- 5 In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein.
- 10 Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

15

Vectors, Host Cells, and Protein Production

- The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.
- 20

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

- 25

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The

- 30

expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or

- 5 UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples

- 10 of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila S2* and *Spodoptera Sf9* cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture media and conditions for the above-described host cells are known in the art.

- 15 Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

- 20 Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods 25 are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

- A polypeptide of this invention can be recovered and purified from 30 recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography,

phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

- 5 Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.
- 10
- 15
- 20 In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., coding sequence), and/or to include genetic material (e.g., heterologous polynucleotide sequences) that is operably associated with the polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g., promoter and/or enhancer) and endogenous polynucleotide sequences via homologous recombination (see, e.g., U.S. Patent No. 5,641,670, issued June 24, 1997; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller et al.,
- 25
- 30

Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra et al., Nature 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

5

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

- 10 The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.
- 15 Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids
- 20 containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides 25 can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include *in situ* hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

- 30 Precise chromosomal location of the polynucleotides can also be achieved using fluorescence *in situ* hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however,

polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques." Pergamon Press, New York (1988).

- For chromosome mapping, the polynucleotides can be used individually (to
- 5 mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.
- 10 Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins
- 15 University Welch Medical Library).) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.
- Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined.
- 20 First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide
- 25 and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.
- Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using
- 30 polynucleotides of the present invention. Any of these alterations (altered expression,

chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods

- 5 rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J.
- 10 Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can
- 15 be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel.

- 25 In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA
- 30 markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using

- 5 this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification
10 techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co.
15 (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

20 There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by
25 organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences
30 in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA

antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

- 5 Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell. Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected *in vivo* by imaging. Antibody labels or markers for *in vivo* imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety

needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein.

- 5 In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which

- 10 involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

- 15 Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to 20 activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

- 25 Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

- 30 At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also

be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

5

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules

- 10 may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

- A polypeptide or polynucleotide of the present invention may be useful in
- 15 treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or
- 20 disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

- A polynucleotide or polypeptide of the present invention may be useful in
- 25 treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes
- 30 include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency,

Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

- 5 Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.
- 10 A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

- 15 Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune inflammatory eye disease.
- 20
- 25
- 30

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

5 A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in
10 this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

15 Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)
20

25 **Hyperproliferative Disorders**

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present
30 invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

- 5 Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to
10 neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

- Similarly, other hyperproliferative disorders can also be treated or detected by
15 a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to:
hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenström's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia,
20 located in an organ system listed above.

Infectious Disease

- A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.
30 Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the

present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes

- 5 Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or
- 10 symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox , hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually
- 15 transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present

- 20 invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Nocardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses,
- 25 Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections (e.g., Actinobacillus, Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can
- 30 cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis,

opportunistic infections (e.g., AIDS related infections), paronychia, prosthetic-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis,

5 Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocytoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

10 Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidioidomycosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and

15 Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

20 Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

25 A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See,

Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis,

- 5 reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vasculature (including vascular and lymphatics), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration 10 occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used

- 15 prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure 20 ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a 25 polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, 30 Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotactic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or

- 5 site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotactic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotactic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

- 10 It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

- 15 A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

- 20 Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable

of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell

5 membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

10 The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

15 Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

20 Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers.

25 The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

30 Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a

- candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity , and (b)
- 5 determining if a biological activity of the polypeptide has been altered.

Other Activities

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as

10 discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be

15 used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, cardiac rhythms, depression (including depressive disorders), tendency for violence,

20 tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated

30 nucleic acid molecule comprising a nucleotide sequence which is at least 95%

identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of

- 5 positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of

- 10 positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the

- 15 range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

- Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

- 25 A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X
30 in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

25 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

30 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X

- 5 wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and
- 10 determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence

- 15 selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

- 20 A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X
- 25 wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

- 30 The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least

one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted

- 5 protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1;
- 10 and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

- 15 Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1
- 20 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

- 25 Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the

- 5 Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

- Further preferred is an isolated polypeptide comprising an amino acid
10 sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

- 15 Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

- 20 Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

- 25 Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

- 30 Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in

the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence
5 at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a
10 polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained
15 in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group
20 consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least
25 one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence
30 selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide

comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA

- 5 clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

- 10 Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is

- 15 any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

- Also preferred is the above method for identifying the species, tissue or cell
20 type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

- 25 Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least
30 one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino

acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

- 5 In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at

- 10 least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number
- 15 shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

- 20 Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

- 25 Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

- 30 Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is

expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of

- 5 SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1
- 10 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an

- 15 isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

20

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector.

- 25 Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being
- 30 isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

	<u>Vector Used to Construct Library</u>	<u>Corresponding Deposited</u>
	<u>Plasmid</u>	
	Lambda Zap	pBluescript (pBS)
	Uni-Zap XR	pBluescript (pBS)
5	Zap Express	pBK
	lafmid BA	plafmid BA
	pSport1	pSport1
	pCMVSport 2.0	pCMVSport 2.0
	pCMVSport 3.0	pCMVSport 3.0
10	pCR®2.1	pCR®2.1
	Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap	
	XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos.	
	5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res.	
	16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res.	
15	17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are	
	commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey	
	Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and	
	pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain	
	XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+	
20	and KS. The S and K refers to the orientation of the polylinker to the T7 and T3	
	primer sequences which flank the polylinker region ("S" is for SacI and "K" is for	
	KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to	
	the orientation of the f1 origin of replication ("ori"), such that in one orientation,	
	single stranded rescue initiated from the f1 ori generates sense strand DNA and in the	
25	other, antisense.	
	Vectors pSport1, pCMVSport 2.0 and pCMVSport 3.0, were obtained from	
	Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors	
	contain an ampicillin resistance gene and may be transformed into E. coli strain	
	DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et	
30	al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University,	
	NY) contains an ampicillin resistance gene and can be transformed into E. coli strain	

XL-1 Blue, Vector pCR®2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into *E. coli* strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al.,

- 5 Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional

- 10 plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample
15 may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ

- 20 ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported.

The oligonucleotide is labeled, for instance, with ^{32}P - γ -ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular

- 25 Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).)
The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above.
The transformants are plated on 1.5% agar plates (containing the appropriate selection
30 agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate.
These plates are screened using Nylon membranes according to routine methods for

bacterial colony screening (e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the

- 5 SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 µl of reaction mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is
- 10 1.5-5 mM MgCl₂, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with
- 15 expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

- Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

- Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

- 30 This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with

phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

5 This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific 10 to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

15 A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

Example 3: Tissue Distribution of Polypeptide

20 Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P³² using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After 25 labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) 30 or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to

manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

5 **Example 4: Chromosomal Mapping of the Polynucleotides**

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is 10 repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in 15 the particular somatic cell hybrid.

Example 5: Bacterial Expression of a Polypeptide

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA 20 sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning 25 sites.

The pQE-9 vector is digested with BamHI and XbaI and the amplified 30 fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain

M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^R).

Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D. ⁶⁰⁰) of between 0.4 and 0.6. IPTG

10 (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic

15 agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The
20 QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

25 The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4,
30 containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250

mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

- In addition to the above expression vector, the present invention further
- 5 includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains:
- 10 1) a neomycinphosphotransferase gene as a selection marker, 2) an *E. coli* origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (*lacIq*). The origin of replication (*oriC*) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

- DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating
- 15 the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

- 20 The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

Example 6: Purification of a Polypeptide from an Inclusion Body

- The following alternative method can be used to purify a polypeptide
- 25 expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

- Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit
- 30 weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM

Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate

- 5 is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the

- 10 pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA

- 15 by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The

- 20 filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

- 25 Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium

acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the
5 above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 μ g of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

10

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector
15 contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak
20 Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such
25 as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., Virology 170:31-39 (1989).

30 Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in

Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in

- 5 Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment

- 10 then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a 15 commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

- The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified 20 by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

- Five µg of a plasmid containing the polynucleotide is co-transfected with 1.0 µg of a commercially available linearized baculovirus DNA ("BaculoGold™ 25 baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One µg of BaculoGold™ virus DNA and 5 µg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are 30 added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded

in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

- 5 After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell
- 10 culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded
- 15 in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

- To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 µCi of ³⁵S-methionine and 5 µCi ³⁵S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation.
- 20 The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).
 - 25

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

- 30 **Example 8: Expression of a Polypeptide in Mammalian Cells**

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a

chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession

- 5 No. 209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the
10 polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

- 15 A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

- 20 The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

- The amplified fragment is then digested with the same restriction enzyme and
25 purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

- Chinese hamster ovary cells lacking an active DHFR gene is used for
30 transfection. Five µg of the expression plasmid pC6 is cotransfected with 0.5 µg of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo

- contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μ M, 2 μ M, 5 μ M, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 μ M. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.
- 15 **Example 9: Protein Fusions**
The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see
20 also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create
25 chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.
30 Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These

primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be

5 destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

10 If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

15 Human IgG Fc region:

GGGATCCGGAGCCAAATCTTCTGACAAAACTCACACATGCCACC GTGC
CCAGCACCTGAATTGAGGGTGCAACCGTCAGTCTCCTCTCCCCAAAAA
CCCAAGGACACCCCTCATGATCTCCGGACTCCTGAGGTACATGCGTGGT
GGTGGACGTAAGCCACGAAGACCTGAGGTCAAGTTCAACTGGTACGTGG

20 ACGCGTGGAGGTGCATAATGCCAAGACAAAGCCGGAGGAGCAGTA
CAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACT
GGCTGAATGGCAAGGAGTACAAGTCAAGGTCTCCAACAAAGCCCTCCCA
ACCCCCATCGAGAAAACCATCTCAAAGCCAAGGGCAGCCCCGAGAAC
CACAGGTGTACACCCCTGCCCATCCGGATGAGCTGACCAAGAACAG

25 GTCAGCCTGACCTGCCCTGGTCAAAGGCTTCTATCCAAGCGACATGCCGT
GGAGTGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCT
CCCGTGTGGACTCCGACGGCTCTTCTCCTACAGCAAGCTACCGTG
GACAAGAGCAGGTGGCAGCAGGGAAACGTCTCTCATGCTCCGTATGCA
TGAGGCTCTGCACAACCACTACACG CAGAAGAGCCTCTCCGTCTCCGG
30 GTAAATGAGTGCACGGCCCGACTCTAGAGGAT (SEQ ID NO:1)

Example 10: Production of an Antibody from a Polypeptide

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

- In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., *Nature* 256:495 (1975); Köhler et al., *Eur. J. Immunol.* 6:511 (1976); Köhler et al., *Eur. J. Immunol.* 6:292 (1976); Hammerling et al., in: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

- The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (*Gastroenterology* 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

- Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is

possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody

- 5 whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulian et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

25

Example 11: Production Of Secreted Protein For High-Throughput Screening Assays

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in

30 Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each

- 5 well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2×10^5 cells/well in .5ml

- 10 DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate.

- 15 With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45
- 20 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

- Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a 12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6
- 30 hours.

- While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl₂ (anhyd); 0.00130 mg/L CuSO₄·5H₂O; 0.050 mg/L of Fe(NO₃)₃·9H₂O; 0.417 mg/L of FeSO₄·7H₂O; 311.80 mg/L of KCl; 28.64 mg/L of MgCl₂; 48.84 mg/L of MgSO₄; 6995.50 mg/L of
- 5 NaCl; 2400.0 mg/L of NaHCO₃; 62.50 mg/L of NaH₂PO₄·H₂O; 71.02 mg/L of Na₂HPO₄; .4320 mg/L of ZnSO₄·7H₂O; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitic Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of
- 10 Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L-Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H₂O; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H₂O; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H₂O; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalanine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tyrosine-2Na-2H₂O; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of
- 15 Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B₁₂; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine; 0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL;
- 20 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L
- 25 DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.
- 30

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

- 5 On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the

- 10 polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

15 **Example 12: Construction of GAS Reporter Construct**

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The 20 binding of a protein to these elements alter the expression of the associated gene.

- GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

- The STATs are activated to translocate from the cytoplasm to the nucleus 30 upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2,

Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-

- 5 51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN- α , IFN- γ , and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS
- 10 motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

- Therefore, activation of the Jaks-STATs pathway, reflected by the binding of
- 15 the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

	<u>Ligand</u>	<u>tyk2</u>	<u>JAKs</u>			<u>STATs</u>	<u>GAS(elements) or ISRE</u>	
			<u>Jak1</u>	<u>Jak2</u>	<u>Jak3</u>			
<u>IFN family</u>								
5	IFN-a/B	+	+	-	-	1,2,3	ISRE	
	IFN-g		+	+	-	1	GAS (IRF1>Lys6>IFP)	
	Il-10	+	?	?	-	1,3		
<u>gp130 family</u>								
10	IL-6 (Pleiotrophic)	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)	
	Il-11(Pleiotrophic)	?	+	?	?	1,3		
	OnM(Pleiotrophic)	?	+	+	?	1,3		
	LIF(Pleiotrophic)	?	+	+	?	1,3		
	CNTF(Pleiotrophic)	-/+	+	+	?	1,3		
15	G-CSF(Pleiotrophic)	?	+	?	?	1,3		
	IL-12(Pleiotrophic)	+	-	+	+	1,3		
<u>g-C family</u>								
20	IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS	
	IL-4 (lymph/myeloid)	-	+	-	+	6	GAS (IRF1 = IFP >> Ly6)(IgH)	
	IL-7 (lymphocytes)	-	+	-	+	5	GAS	
	IL-9 (lymphocytes)	-	+	-	+	5	GAS	
	IL-13 (lymphocyte)	-	+	?	?	6	GAS	
	IL-15	?	+	?	+	5	GAS	
25	<u>gp140 family</u>							
	IL-3 (myeloid)	-	-	+	-	5	GAS (IRF1>IFP>>Ly6)	
	IL-5 (myeloid)	-	-	+	-	5	GAS	
	GM-CSF (myeloid)	-	-	+	-	5	GAS	
30	<u>Growth hormone family</u>							
	GH	?	-	+	-	5		
	PRL	?	+/-	+	-	1,3,5		
	EPO	?	-	+	-	5	GAS(B-CAS>IRF1=IFP>>Ly6)	
35	<u>Receptor Tyrosine Kinases</u>							
	EGF	?	+	+	-	1,3	GAS (IRF1)	
	PDGF	?	+	+	-	1,3		
	CSF-1	?	+	+	-	1,3	GAS (not IRF1)	

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., *Immunity* 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5'

- 5 primer is:
- 10 5':GCGCCTCGAGATTCCCCGAAATCTAGATTCCCCGAAATGATTCCCC
GAAATGATTCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTGCAAAGCCTAGGC:3'
(SEQ ID NO:4)

- 15 PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

- 20 5':CTCGAGATTCCCCGAAATCTAGATTCCCCGAAATGATTCCCCGAAA
TGATTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCG
CCCCTACTCCGCCCATCCCGCCCTACTCCGCCCAGTCCGCCATTTCT
CCGCCCCATGGCTGACTAATTTTTATTATGCAGAGGCCAGGCC
TCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTGGAGGCCT
25 AGGTTTGCAAAAGCTT:3' (SEQ ID NO:5)

- With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol

acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and

- 5 XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-

- 10 SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using Sall and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding
15 as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described

- 20 in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

25

Example 13: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12.

- 30 Thus, factors that increase SEAP activity indicate the ability to activate the Jak-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells

(ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-

- 5 SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml genticin selected.
- 10 Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1%Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life

- 15 Technologies) with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

- 20 During the incubation period, count cell concentration, spin down the required number of cells (10^7 per transfection), and resuspend in OPTI-MEM to a final concentration of 10^7 cells/ml. Then add 1ml of 1×10^7 cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

- 25 The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

- 30 On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100, 000 cells per well).

- After all the plates have been seeded, 50 ul of the supernatants are transferred
- 5 directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

- The 96 well dishes containing Jurkat cells treated with supernatants are placed
- 10 in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source
- 15 of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

- The above protocol may be used in the generation of both transient, as well as,
- 20 stable transfected cells, which would be apparent to those of skill in the art.

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

- The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate
- 25 myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATs signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.
- 30 To transiently transfet U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell

Growth & Differentiation, 5:259-265) is used. First, harvest 2×10^7 U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

5 Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 1 mM MgCl_2 , and 675 uM CaCl_2 . Incubate at 37°C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then
10 resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting 1×10^8 cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of 5×10^5 cells/ml. Plate 200 ul cells per well in the 96-well plate (or 1×10^5 cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon
20 gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

25 When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor).

- 5 The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 10:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)
5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

- Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes Xhol/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.
- 15

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

- PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.
- 25

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine

- 30

growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS

- 5 (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as 5×10^5

- 10 cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to 1×10^5 cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold

- 15 induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

Example 16: High-Throughput Screening Assay for T-cell Activity

- NF- κ B (Nuclear Factor κ B) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF- κ B regulates the expression of genes involved in immune cell activation, control of apoptosis (NF- κ B appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

- In non-stimulated conditions, NF- κ B is retained in the cytoplasm with I- κ B (Inhibitor κ B). However, upon stimulation, I- κ B is phosphorylated and degraded, causing NF- κ B to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- κ B include IL-2, IL-6, GM-CSF, ICAM-1 and class 1 MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF- κ B promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF- κ B would be useful in treating diseases. For example, inhibitors of NF- κ B could be used to treat those

- 5 diseases related to the acute or chronic activation of NF- κ B, such as rheumatoid arthritis.

To construct a vector containing the NF- κ B promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF- κ B binding site (GGGGACTTCCCC) (SEQ ID NO:8), 18 bp of sequence complementary

- 10 to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:
5':GCGGCCTCGAGGGGACTTCCCCGGGGACTTCCGGGGACTTCCGGGAC
TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

- 15 5':GCGGCAAGCTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene)

- 20 Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

- 25 5':CTCGAGGGGACTTCCCCGGGGACTTCCGGGGACTTCCGGGACTTTCC
ATCTGCCATCTCAATTAGTCAGCAACCATACTCCGCCCTAACTCCGCC
ATCCCGCCCTAACTCCGCCAGTCCGCCATTCTCCGCCCATGGCTGA
CTAATTTTTTATTATGCAGAGGCCGAGGCCCTCGGCCTGTGAGCTA
TTCCAGAAGTAGTGAGGAGGCTTTTGAGGCCAGGCTTTGCAAAAAA
GCTT:3' (SEQ ID NO:10)

- 30 Next, replace the SV40 minimal promoter element present in the pSEAP2- promoter plasmid (Clontech) with this NF- κ B/SV40 fragment using XhoI and

HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF- κ B/SV40/SEAP cassette is removed from the above NF- κ B/SEAP vector using restriction enzymes

- 5 SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF- κ B/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF- κ B/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly,

- 10 the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

- 15 As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 μ l of 2.5x dilution buffer into Optiplates containing 35 μ l of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

- 20 Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 μ l Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 μ l Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.
- 25

30 Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Buffer Formulation:

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

5 Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants

which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

- 5 The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-4 (Molecular Probes, Inc.; catalog no. F-14202), used here.
- 10 For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO₂ incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.
A stock solution of 1 mg/ml fluo-4 is made in 10% pluronic acid DMSO. To
- 15 load the cells with fluo-4 , 50 ul of 12 ug/ml fluo-4 is added to each well. The plate is incubated at 37°C in a CO₂ incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.
For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10⁶ cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml
- 20 fluo-4 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10⁶ cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.
- 25 For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-4 . The supernatant is added to the well, and a change in fluorescence is detected.
To measure the fluorescence of intracellular calcium, the FLIPR is set for the
- 30 following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm;

and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca⁺⁺ concentration.

5 **Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity**

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St.

Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford,MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar

- 5 Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford,MA) are used to cover the Loprodynne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

- To prepare extracts, A431 cells are seeded onto the nylon membranes of
10 Loprodynne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na₃VO₄, 2 mM Na₄P₂O₇ and a cocktail of protease inhibitors (# 1836170) obtained from Boheringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well
15 catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000 x g.

- Test the filtered extracts for levels of tyrosine kinase activity. Although many
20 methods of detecting tyrosine kinase activity are known, one method is described here.

- Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include
25 PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and

PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg₂₊ (5mM

- 5 ATP/50mM MgCl₂), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl₂, 5 mM MnCl₂, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30°C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

- 10 The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mm EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavidin coated 96 well plate to associate with the biotinylated peptide.

- 15 Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

- 20 Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

- 25 **Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity**

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be

- 30 used. For example, as described below one particular assay can detect tyrosine

phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by

- 5 substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G

plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against

- 10 Erk-1

and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

- 15 A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

- 20 After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place

of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the

- 25 Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from

5 these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

10 PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTHERM Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct 15 sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

20 Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenin-deoxy-uridine 5'-triphosphate (Boehringer Manheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe 25 is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology,

30 Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv.

et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and 5 translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

10 A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

15 For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10. The wells are blocked so that non-specific binding of the polypeptide to 20 the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

25 Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

30 Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard

curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

5

Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the 10 secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 15 $\mu\text{g}/\text{kg}/\text{day}$ to 10 $\text{mg}/\text{kg}/\text{day}$ of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 $\text{mg}/\text{kg}/\text{day}$, and most preferably for humans between about 0.01 and 1 $\text{mg}/\text{kg}/\text{day}$ for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 $\mu\text{g}/\text{kg}/\text{hour}$ to about 50 $\mu\text{g}/\text{kg}/\text{hour}$, either by 20 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention 25 are administered orally, rectally, parenterally, intracistemally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to 30 modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or mirocapsules. Sustained-release matrices include polylactides (U.S. Pat. No.

- 5 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D-(-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions
10 also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP
15 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms)
unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

- For parenteral administration, in one embodiment, the secreted polypeptide is
20 formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other
25 compounds that are known to be deleterious to polypeptides.

- Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood
30 of the recipient. Examples of such carrier vehicles include water, saline, Ringer's

solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to

- 5 recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as
- 10 polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

- 15 The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile.

- 20 Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

- 25 Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

- 30 The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical

compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the

- 5 polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 24: Method of Treating Decreased Levels of the Polypeptide

- It will be appreciated that conditions caused by a decrease in the standard or
- 10 normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to
 - 15 increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

20

Example 25: Method of Treating Increased Levels of the Polypeptide

- Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as
- 25 cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense

- 30 polynucleotide is provided in Example 23.

Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and

- 5 separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS,
- 10 penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

- 15 pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

- 20 The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.
- 25

- 30 The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector.

The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media,

- 5 containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media.
- If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or

- 15 after having been grown to confluence on cytodex 3 microcarrier beads.

Example 27: Method of Treatment Using Gene Therapy - In Vivo

- 20 Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide. The polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO90/11092, WO98/11779; U.S. Patent NO. 5693622, 5705151, 5580859; Tabata H. et al. (1997) Cardiovasc. Res. 35(3):470-479, Chao J et al. (1997) Pharmacol. Res. 35(6):517-522, Wolff J.A. (1997) Neuromuscul. Disord.
- 25 7(5):314-318, Schwartz B. et al. (1996) Gene Ther. 3(5):405-411, Tsurumi Y. et al. (1996) Circulation 94(12):3281-3290 (incorporated herein by reference).
- 30

The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

5 The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present
10 invention may also be delivered in liposome formulations (such as those taught in Felgner P.L. et al. (1995) Ann. NY Acad. Sci. 772:126-139 and Abdallah B. et al. (1995) Biol. Cell 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

15 The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies
20 have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

25 The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It
30 is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection

into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In*

- 5 *vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

20 The dose response effects of injected polynucleotide in muscle *in vivo* is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

25 Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that

5 quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice. The results of the above experimentation in mice can be used to extrapolate proper dosages and other treatment parameters in humans and other animals using naked

10 DNA.

Example 28: Transgenic Animals.

The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e.g., baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

20 Any technique known in the art may be used to introduce the transgene (i.e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., Appl. Microbiol. Biotechnol. 40:691-698 (1994); Carver et al., Biotechnology (NY) 11:1263-1270 (1993); Wright et al., Biotechnology 15 (NY) 9:830-834 (1991); and Hoppe et al., U.S. Pat. No. 4,873,191 (1989)); retrovirus mediated gene transfer into germ lines (Van der Putten et al., Proc. Natl. Acad. Sci., USA 82:6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., Cell 56:313-321 (1989)); electroporation of cells or embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814 (1983)); introduction of the 25 polynucleotides of the invention using a gene gun (see, e.g., Ulmer et al., Science 259:1745 (1993); introducing nucleic acid constructs into embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-

mediated gene transfer (Lavitrano et al., Cell 57:717-723 (1989); etc. For a review of such techniques, see Gordon, "Transgenic Animals," Intl. Rev. Cytol. 115:171-229 (1989), which is incorporated by reference herein in its entirety.

Any technique known in the art may be used to produce transgenic clones
5 containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell et al., Nature 380:64-66 (1996); Wilmut et al., Nature 385:810-813 (1997)).

The present invention provides for transgenic animals that carry the transgene
10 in all their cells, as well as animals which carry the transgene in some, but not all their cells, *i.e.*, mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, *e.g.*, head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et
15 al. (Lasko et al., Proc. Natl. Acad. Sci. USA 89:6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site
20 of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only
25 that cell type, by following, for example, the teaching of Gu et al. (Gu et al., Science 265:103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the
30 recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA

expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and reverse transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also

- 5 be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more

- 10 than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for
15 screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

25 **Example 29: Knock-Out Animals.**

Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (E.g., see Smithies et al., Nature 317:230-234 (1985); Thomas & Capecchi, Cell 51:503-512 (1987); Thompson et al., Cell 5:313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding

regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfet cells that express polypeptides of the invention *in vivo*. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (e.g., see Thomas & Capecchi 1987 and Thompson 1989, *supra*). However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors that will be apparent to those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (e.g., knockouts) are administered to a patient *in vivo*. Such cells may be obtained from the patient (i.e., animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (e.g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered *in vitro* using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e.g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc. The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, e.g., in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e.g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U.S. Patent No. 5,399,349; and

- 5 Mulligan & Wilson, U.S. Patent No. 5,460,959 each of which is incorporated by reference herein in its entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For
10 example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

- Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological
15 function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

- It will be clear that the invention may be practiced otherwise than as
20 particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

- The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other
25 disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>178</u> line <u>N/A</u>														
B. IDENTIFICATION OF DEPOSIT		Further deposits are identified on an additional sheet <input type="checkbox"/>												
Name of depositary institution American Type Culture Collection														
Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America														
Date of deposit <u>May 18, 1998</u>	Accession Number <u>209878</u>													
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)		This information is continued on an additional sheet <input type="checkbox"/>												
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)														
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)														
<table border="1"> <tr> <td colspan="2">For receiving Office use only</td> </tr> <tr> <td><input checked="" type="checkbox"/> This sheet was received with the international application</td> <td></td> </tr> <tr> <td colspan="2">Authorized officer <u>Misty Walker</u> <u>International Division</u> <u>703-205-2032</u> <u>mwalker@uspto.gov</u></td> </tr> <tr> <td colspan="2">For International Bureau use only</td> </tr> <tr> <td colspan="2"><input type="checkbox"/> This sheet was received by the International Bureau on:</td> </tr> <tr> <td colspan="2">Authorized officer</td> </tr> </table>			For receiving Office use only		<input checked="" type="checkbox"/> This sheet was received with the international application		Authorized officer <u>Misty Walker</u> <u>International Division</u> <u>703-205-2032</u> <u>mwalker@uspto.gov</u>		For International Bureau use only		<input type="checkbox"/> This sheet was received by the International Bureau on:		Authorized officer	
For receiving Office use only														
<input checked="" type="checkbox"/> This sheet was received with the international application														
Authorized officer <u>Misty Walker</u> <u>International Division</u> <u>703-205-2032</u> <u>mwalker@uspto.gov</u>														
For International Bureau use only														
<input type="checkbox"/> This sheet was received by the International Bureau on:														
Authorized officer														

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>179</u> , line <u>N/A</u>		
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>		
Name of depositary institution <u>American Type Culture Collection</u>		
Address of depositary institution (<i>including postal code and country</i>) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>		
Date of deposit <u>February 26, 1997</u>	Accession Number	<u>97898</u>
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>) This information is continued on an additional sheet <input type="checkbox"/>		
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)		
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)		
For receiving Office use only		
<input checked="" type="checkbox"/> This sheet was received with the international application		
For International Bureau use only		
<input type="checkbox"/> This sheet was received by the International Bureau on:		
Authorized officer <u>Mary Walker</u> <u>Examination Division</u> <u>703-308-3032</u> <u>Int'l Search and Exam Div</u>	Authorized officer	

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>179</u> , line <u>N/A</u>		
B. IDENTIFICATION OF DEPOSIT		Further deposits are identified on an additional sheet <input type="checkbox"/>
Name of depository institution <u>American Type Culture Collection</u>		
Address of depository institution (<i>including postal code and country</i>) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>		
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209044</u>	
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)		This information is continued on an additional sheet <input type="checkbox"/>
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)		
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g. "Accession Number of Deposit"</i>)		
For receiving Office use only		For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer <i>[Signature]</i> [Signature] [Signature]		Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>179</u> , line <u>N/A</u>		
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>		
Name of depositary institution American Type Culture Collection		
Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America		
Date of deposit August 28, 1997	Accession Number 209225	
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>) This information is continued on an additional sheet <input type="checkbox"/>		
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)		
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)		
For receiving Office use only		For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer <i>[Signature]</i> Microbial Division 703-270-3202		Authorized officer

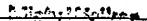
INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>180</u> , line <u>N/A</u>								
B. IDENTIFICATION OF DEPOSIT		Further deposits are identified on an additional sheet <input type="checkbox"/>						
Name of depositary institution American Type Culture Collection								
Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America								
Date of deposit <u>March 7, 1997</u>	Accession Number	<u>97923</u>						
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)		This information is continued on an additional sheet <input type="checkbox"/>						
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)								
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>)								
The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)								
<table border="0"> <tr> <td style="text-align: center;">For receiving Office use only</td> <td style="text-align: center;">For International Bureau use only</td> </tr> <tr> <td><input checked="" type="checkbox"/> This sheet was received with the international application</td> <td><input type="checkbox"/> This sheet was received by the International Bureau on:</td> </tr> <tr> <td>Authorized officer Kathy Walker International Division 703-305-2000</td> <td>Authorized officer</td> </tr> </table>			For receiving Office use only	For International Bureau use only	<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:	Authorized officer Kathy Walker International Division 703-305-2000	Authorized officer
For receiving Office use only	For International Bureau use only							
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:							
Authorized officer Kathy Walker International Division 703-305-2000	Authorized officer							

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>180</u> , line <u>N/A</u>		
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>		
Name of depositary institution American Type Culture Collection		
Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America		
Date of deposit May 22, 1997	Accession Number 209071	
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>) This information is continued on an additional sheet <input type="checkbox"/>		
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)		
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)		
For receiving Office use only		For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer	 Jennifer L. Johnson International Division 703-235-7232 jennifer.l.johnson@uspto.gov	
Authorized officer		

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>180</u> , line <u>N/A</u>											
B. IDENTIFICATION OF DEPOSIT		Further deposits are identified on an additional sheet <input type="checkbox"/>									
Name of depositary institution American Type Culture Collection											
Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America											
Date of deposit <u>February 12, 1998</u>	Accession Number <u>209626</u>										
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)		This information is continued on an additional sheet <input type="checkbox"/>									
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)											
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)											
<table border="1"> <tr> <td colspan="2">For receiving Office use only</td> <td>For International Bureau use only</td> </tr> <tr> <td><input checked="" type="checkbox"/> This sheet was received with the international application</td> <td colspan="2"><input type="checkbox"/> This sheet was received by the International Bureau on:</td> </tr> <tr> <td colspan="2">Authorized officer: <u>J. M. Johnson</u> <u>International Division</u> <u>709-CCE-8680</u></td> <td>Authorized officer</td> </tr> </table>			For receiving Office use only		For International Bureau use only	<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:		Authorized officer: <u>J. M. Johnson</u> <u>International Division</u> <u>709-CCE-8680</u>		Authorized officer
For receiving Office use only		For International Bureau use only									
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:										
Authorized officer: <u>J. M. Johnson</u> <u>International Division</u> <u>709-CCE-8680</u>		Authorized officer									

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>181</u> , line <u>N/A</u>		
B. IDENTIFICATION OF DEPOSIT		Further deposits are identified on an additional sheet <input type="checkbox"/>
Name of depositary institution American Type Culture Collection		
Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America		
Date of deposit March 20, 1998	Accession Number 209683	
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>) This information is continued on an additional sheet <input type="checkbox"/>		
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)		
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)		
For receiving Office use only		For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer <i>[Signature]</i> U.S. Patent and Trademark Office 703-305-9000		Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>181</u> , line <u>N/A</u>		
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>		
Name of depositary institution American Type Culture Collection		
Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America		
Date of deposit February 25, 1998	Accession Number 209641	
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>) This information is continued on an additional sheet <input type="checkbox"/>		
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)		
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)		
For receiving Office use only		For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer <i>[Signature]</i> FBI - D.C. - C.R.P. FBI Laboratory		Authorized officer

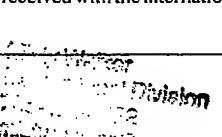
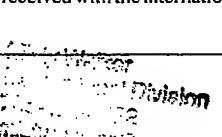
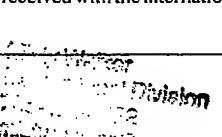
INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>181</u> , line <u>N/A</u>		
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>		
Name of depositary institution American Type Culture Collection		
Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America		
Date of deposit <u>July 9, 1997</u>	Accession Number <u>209141</u>	
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>) This information is continued on an additional sheet <input type="checkbox"/>		
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)		
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)		
For receiving Office use only		For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer <i>[Signature]</i>		Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>181</u> , line <u>N/A</u>								
B. IDENTIFICATION OF DEPOSIT		Further deposits are identified on an additional sheet <input type="checkbox"/>						
Name of depository institution American Type Culture Collection								
Address of depository institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America								
Date of deposit February 26, 1997	Accession Number 97901							
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)		This information is continued on an additional sheet <input type="checkbox"/>						
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)								
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>)								
The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g. "Accession Number of Deposit"</i>)								
<table border="0"> <tr> <td style="text-align: center;">For receiving Office use only</td> <td style="text-align: center;">For International Bureau use only</td> </tr> <tr> <td><input checked="" type="checkbox"/> This sheet was received with the international application</td> <td><input type="checkbox"/> This sheet was received by the International Bureau on:</td> </tr> <tr> <td>Authorized officer </td> <td>Authorized officer</td> </tr> </table>			For receiving Office use only	For International Bureau use only	<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:	Authorized officer 	Authorized officer
For receiving Office use only	For International Bureau use only							
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:							
Authorized officer 	Authorized officer							

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>181</u> , line <u>N/A</u>		
B. IDENTIFICATION OF DEPOSIT		Further deposits are identified on an additional sheet <input type="checkbox"/>
Name of depositary institution <u>American Type Culture Collection</u>		
Address of depositary institution (<i>including postal code and country</i>) <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>		
Date of deposit <u>May 15, 1997</u>	Accession Number	<u>209047</u>
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)		This information is continued on an additional sheet <input type="checkbox"/>
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)		
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)		
For receiving Office use only		For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer <u>[Signature]</u>		Authorized officer

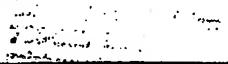
INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>181</u> , line <u>N/A</u>		
B. IDENTIFICATION OF DEPOSIT		Further deposits are identified on an additional sheet <input type="checkbox"/>
Name of depositary institution American Type Culture Collection		
Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America		
Date of deposit	Accession Number	
May 18, 1998	209877	
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)		This information is continued on an additional sheet <input type="checkbox"/>
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)		
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>)		
The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)		
For receiving Office use only		For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer	Sally Walker INTERNATIONAL DIVISION 703-605-0802	
Authorized officer		

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>182</u> , line <u>N/A</u>		
B. IDENTIFICATION OF DEPOSIT		Further deposits are identified on an additional sheet <input type="checkbox"/>
Name of depositary institution American Type Culture Collection		
Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America		
Date of deposit January 14, 1998	Accession Number 209580	
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)		This information is continued on an additional sheet <input type="checkbox"/>
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)		
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)		
For receiving Office use only <input checked="" type="checkbox"/> This sheet was received with the international application		For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 		Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<p>A. The indications made below relate to the microorganism referred to in the description on page <u>185</u>, line <u>N/A</u></p>		
<p>B. IDENTIFICATION OF DEPOSIT</p>		<p>Further deposits are identified on an additional sheet <input type="checkbox"/></p>
<p>Name of depositary institution American Type Culture Collection</p>		
<p>Address of depositary institution (<i>including postal code and country</i>) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America</p>		
Date of deposit		Accession Number
<u>May 22, 1998</u>		<u>209889</u>
<p>C. ADDITIONAL INDICATIONS (leave blank if not applicable)</p>		<p>This information is continued on an additional sheet <input type="checkbox"/></p>
<p>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)</p>		
<p>E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)</p> <p>The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications e.g., "Accession Number of Deposit"</i>)</p>		
<p>For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p>		<p>For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p>
<p>Authorized officer</p> <p><i>[Signature]</i> YOLANDA S. BROWN 7/22/98 <i>[Signature]</i></p>		<p>Authorized officer</p> <p><i>[Signature]</i></p>

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
 - 5 (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- 10 (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X,
- 15 (f) a polynucleotide which is a variant of SEQ ID NO:X;
- (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
- (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
- 20 (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.
- 25

2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.

5 3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

10 4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

15 5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

20 7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

25 8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

9. A recombinant host cell produced by the method of claim 8.

30 10. The recombinant host cell of claim 9 comprising vector sequences.

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

5 (b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;

(c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

10 (d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

15 (g) a variant of SEQ ID NO:Y;

(h) an allelic variant of SEQ ID NO:Y; or

(i) a species homologue of the SEQ ID NO:Y.

12. The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.

25 14. A recombinant host cell that expresses the isolated polypeptide of claim 11.

15. A method of making an isolated polypeptide comprising:
30 (a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and
(b) recovering said polypeptide.

16. The polypeptide produced by claim 15.

17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount
5 of the polypeptide of claim 11 or the polynucleotide of claim 1.

18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

- 10 (a) determining the presence or absence of a mutation in the polynucleotide of
claim 1; and
(b) diagnosing a pathological condition or a susceptibility to a pathological
condition based on the presence or absence of said mutation.

19. A method of diagnosing a pathological condition or a susceptibility to
15 a pathological condition in a subject comprising:

- (a) determining the presence or amount of expression of the polypeptide of
claim 11 in a biological sample; and
(b) diagnosing a pathological condition or a susceptibility to a pathological
condition based on the presence or amount of expression of the polypeptide.

20

20. A method for identifying a binding partner to the polypeptide of claim
11 comprising:

- 25 (a) contacting the polypeptide of claim 11 with a binding partner; and
(b) determining whether the binding partner effects an activity of the
polypeptide.

21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

22. A method of identifying an activity in a biological assay, wherein the
30 method comprises:

- (a) expressing SEQ ID NO:X in a cell;
(b) isolating the supernatant;

303

- (c) detecting an activity in a biological assay; and
- (d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 20.

1

Sequence Listing

<110> Human Genome Sciences, Inc., et al.

<120> 71 Human Secreted Proteins

<130> PZ030PCT

<140> Unassigned

<141> 1999-07-14

<150> 60/092,956

<151> 1998-07-15

<150> 60/092,921

<151> 1998-07-15

<150> 60/092,922

<151> 1998-07-15

<160> 262

<170> PatentIn Ver. 2.0

<210> 1

<211> 733

<212> DNA

<213> Homo sapiens

<400> 1

gggatccgga	gcccaaatct	tctgacaaaa	ctcacacatg	cccaccgtgc	ccagcacctg	60
aattcgaggg	tgcaccgtca	gtcttcctct	tccccccaaa	acccaaggac	accctcatga	120
tctcccgac	tcctgaggc	acatgcgtgg	tggtggacgt	aagccacgaa	gaccctgagg	180
tcaagtcaa	ctggtaacgtg	gacggcgtgg	aggtgcataa	tgccaagaca	aagccgcggg	240
aggagcagta	caacagcacg	taccgtgtgg	tcagcgtcct	caccgtcctg	caccaggact	300
ggctgaatgg	caaggaggtac	aagtgcagg	tctccaacaa	agccctccca	accccccattcg	360
agaaaaacat	ctccaaagcc	aaagggcagc	cccgagaacc	acaggtgtac	accctgcccc	420
catcccgaaa	tgagctgacc	aagaaccagg	tcagcctgac	ctgcctggtc	aaaggcttct	480
atccaagcga	catcgccgtg	gagtgggaga	gcaatgggca	gcccggagaac	aactacaaga	540
ccacgcctcc	cgtgctggac	tccgacggct	ccttcttcct	ctacagcaag	ctcaccgtgg	600
acaagagcag	gtggcagcag	ggaaacgtct	tctcatgctc	cgtgatgcat	gaggctctgc	660
acaaccacta	cacgcagaag	agccctctccc	tgtctccggg	taaatgagtg	cgacggccgc	720
gactctagag	gat					733

<210> 2

<211> 5

<212> PRT

<213> Homo sapiens

<220>

<221> Site

<222> (3)

<223> Xaa equals any of the twenty naturally occurring L-amino acids

<400> 2

Trp Ser Xaa Trp Ser

1

5

<210> 3
<211> 86
<212> DNA
<213> Homo sapiens

<400> 3
gcccctcgag atttccccga aatctagatt tcccccggaaat gattttcccg aaatgatttc 60
cccgaaaatat ctgccatctc aattag 86

<210> 4
<211> 27
<212> DNA
<213> Homo sapiens

<400> 4
gcggcaagct ttttgcaaag cctaggc 27

<210> 5
<211> 271
<212> DNA
<213> Homo sapiens

<400> 5
ctcgagattt ccccgaaatc tagatttccc cgaaatgatt tcccccggaaat gattttcccg 60
aaatatctgc catctcaatt agtcagcaac catagtcccg ccctctaactc cgcccatccc 120
gccccctaact ccgcccagtt ccgcccattc tccgccccat ggctgactaa ttttttttat 180
ttatgcagag gccgaggccg cctcggccctc tgagctattc cagaagtagt gaggaggctt 240
ttttggaggc cttaggctttt gaaaaaagct t 271

<210> 6
<211> 32
<212> DNA
<213> Homo sapiens

<400> 6
gcgctcgagg gatgacagcg atagaacccc gg 32

<210> 7
<211> 31
<212> DNA
<213> Homo sapiens

<400> 7
gcgaagcttc gcgactcccc ggatccgcct c 31

<210> 8
<211> 12
<212> DNA
<213> Homo sapiens

<400> 8
ggggactttc cc 12

<210> 9
 <211> 73
 <212> DNA
 <213> Homo sapiens

<400> 9
 gcggcctcg a gggactttc ccggggactt tccgggact ttccggact ttccatcctg 60
 ccatctcaat tag 73

<210> 10
 <211> 256
 <212> DNA
 <213> Homo sapiens

<400> 10
 ctcgagggga ct tcccgg gacttccgg ggacttccg ggacttcca tctgccatct 60
 caatttagtca gcaaccatag tcccgccct aactccgccc atccggcccc taactccgccc 120
 cagtccgc cattccgc cccatggctg actaattttt ttatattatg cagaggccga 180
 ggccgcctcg gcctctgagc tattccagaa gtagtgagga ggctttttt gaggcctagg 240
 ct ttttgc aaaa aagctt 256

<210> 11
 <211> 1113
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (393)
 <223> n equals a,t,g, or c

<400> 11
 gatgtccctt tagcttggag gagtttgtta ttaccaccc tctgaaggct acttctgtca 60
 attcatccaa ctcattctca gtccagttt gtttccctgc tggaggag ttgtgatctt 120
 ttggaggaga agaggcattc tggtttttgg aatttttagc catttgctc tggttcttc 180
 ccatctttgt ggatttatct acctttcatc ttcaatgtta gtgacatgtg gatgggtct 240
 ctgagtggtt gtccttcc tttctgttt twagttttt tttaacagt tagccccctt 300
 gctgttaggtc tgctggaktt tgctggatg ccactccaga ccctgtttgc ctgggtatca 360
 ccagtgaggagg ctgcagaaca gcaaagatgg ctncctgttcc tttctctgg aagcttcgtc 420
 tcagagggca cctgccagat gccagccaga gctctctgtt atgaggtgtc tttggccca 480
 tactgggaga ttcccccag tcaggataca aggaggatcg ggacctactt gaggaggcag 540
 tctgaccctt agcagagggtt gaacactgtg ctagggatgc ctctgcttt ttcagagctg 600
 tcaggcgggg cgataaagtgc tgctgaatct gtgtccgcag ccaccccttc ccccagggtgc 660
 tctgtcccaag ggagatgggg gttttattt taagtccca actggggctg ctgcctttt 720
 ttcagagatg ccctgcccag agaggagaaa tctggcagtc tggcctcaga ggccttgc 780
 agctggcgtt ggctccaccc agttcaact tcccaagggtt cttgtttat actgtaaagg 840
 gaaaaaccgc tactcgagcc tcatcaatgg cagacacccc tcccgccgc aagcttaagt 900
 gtcccaggtt gatctcagac tgctgctgtc tggcgtgtga gaatttcaag ccagtggatc 960
 ttagtttgc tgctctgtg ggggtgggac ccattgaacc agactactcg gctccctggc 1020
 ttcagcccccc ttccaggag agtgaagggt tctgtctcat tggcattcca ggagtcaactg 1080
 gtgtatggaa aaaaaaaaaaa aaaggcgcgc cgc 1113

<210> 12

<211> 983
<212> DNA
<213> Homo sapiens

<400> 12

ggcacgaggg cagctgcaga	gctccagggtt tctctgccc	caagggcagg ggctgccc	60
cgcaggat gactctgcct	tccagagcct tggccccc	gggggtggga gtgtgggg	120
tgcataagggtt aaatcaggtc	acagtaagtt gtggggcag	caggtggagc agcagagtgg	180
cactgggagc ttctcttgg	gtgtcggtg tggccttgg	tctgcagcca tcaggtggg	240
gcttggact gacttctct	tctgaaggat gctggaaagg	ttagactggct ttggcagtgc	300
tttagagctcc ggggggttcc	ccctcctaga acatgcaagc	tctcacaccg gtgcgtcatc	360
atcacaccca tcatcaagcc	cacagtggta tactgaaac	ctgccccaca aagacgggtgg	420
actgtctca ggggacccc	atgaaccacc gatggttaca	actatccaat gcctgatggc	480
agacagccag gccaacctcg	gcttccactc tctcttctc	accatcataat cagccaaagt	540
gacctgagtc atgtatgt	aaggttctt ctgcttctc	ttttgtgtc ttttgctgtt	600
tcttcgtccc catactttgt	taactccatg agttaaatgc	tacccatttt cccagacaag	660
tgctgcttct gcaaggaaac	ccttctgtat ccccccaccta	tctgaaaagt acctctccag	720
cttgcattttt caggggtctg	agcgttccctt cccagctgt	catcaccttc ctccatacgc	780
tatggtgtgt tcttgttcc	tctgtcttg tcctttttt	tctgttagat tgtagctct	840
tgctgacagg aaccacgct	gctccagctt catactccc	actgtacag cacagaacct	900
gcttctcaga cttagcaca	atgtttgtt gctgaatgaa	ttaattaaag ataaagcaa	960
aaaaaaaaaaaa aaaaaaaaaactc	gag		983

<210> 13
<211> 973
<212> DNA
<213> Homo sapiens

<400> 13

ggcacgagcc cagcggaaagc	caagccacca ggccccccag	cgtccacgcg gagcatgaac	60
attgaggatg ggcgtgccc	gcccgtccccc gtcccccc	ctggcccccgtt gtaggatgtc	120
ctggcccccac ggggcattgc	tcttcctctg gcttcttctc	ccacccctgg gggccgggtgg	180
agggtggagtgc gccgtgacgt	ctggcccccgg agggggctcc	ccgccccca cctccctgcc	240
cgtggctgc tcttcgcagca	accaggccag ccgggtgatc	tgacacacgga gagacctggc	300
cgagggtccca gccagcatcc	cggtaacac gcggtacctg	aacctgcaag agaacggcat	360
ccagggtgatc cgagcggaca	cgttcaagca cctggggcac	ctggagatcc tgcaagctgag	420
caagaacctg gtgcgcaga	tcgagggtgg cgccttcaac	gggctgccc caacac	480
gctggagatc ttgtacacc	ggctgaccac ggtgcccac	caggccctcg agtacactgtc	540
caagactgcgg gagctctggc	tgcggaaacaa ccccatcgag	agcateccct cctacgcctt	600
caacccgcgtg ccttcgcgtc	ggcgcttgcg cctggcggag	cttaagcggc tggaaatacat	660
ctcgaggcg gccttcgagg	ggctgtcaaa cctgcgtac	ctcaacccctt gcatgtgcaa	720
cctcaaggac atcccaacc	tgacggccctt ggtgcgcctg	gaggagctgg agctgtcggg	780
caacccgcgtg gacctgatcc	gccccggctc cttccagggt	ctcaccagcc tgcgcaagct	840
gtggctcatg cacgcccagg	tagccaccat cgagcgcaac	gccttcgacg acctcaagt	900
gctggaggag ctcaacctgt	cccacaacaa cctgatgtcg	ctgccccacg accttcttac	960
gccccctgcac cgc			973

<210> 14
<211> 1458
<212> DNA
<213> Homo sapiens

<400> 14

ccacgcgtcc gggattttc	aaaagatcca aacagagact	tcctgcatct tctgcctt	60
caacagaagc ggtgatcg	taatgtatgag cctgtggctt	cctttgtgca tttgagcatg	120
ctgttaattaa gatgagatca	gtttctttaga aaaagctt	tcgatccct ctgacgttgc	180

ctgggatctt	tctgttgatt	cgtctttctt	ggagatggg	acagagcatc	tgtggtccag	240
ggaaggtagt	cctctggcct	caattctgtt	gtggatgtgc	agtgataagc	gggcattgcg	300
tgcctcgaaa	gatgcctagt	tcgtggcttc	ctggctgttt	tgtccttctg	tgtctttag	360
ctgttaggggt	ccagctcagg	gagtgggtg	ttggcggcg	ttccgcgggt	ggccctcttg	420
ctttggcga	cctccagggtt	ctgggcatga	gaggccgtgg	cctcattttct	ggtgataac	480
cttttttagtt	taatagcatc	ttaattaga	tcacagcatt	gaattaaaaaa	tttcttctgc	540
aaagaaaaagt	gtggggcata	agacaccggg	aatgagggag	gaggaagaca	gttgtgttt	600
ctctttaaac	cttgagctct	agccgatgca	tttgtcagga	aatacagcac	tttgtcttaa	660
gaaaacaagg	aaggaggccg	ggcgcagttg	ctcacgcctg	taatcccagc	actttggag	720
gcccaggcg	gccccatcacc	tgaggtgggg	agtatagagac	caccctgact	aacatggaga	780
gaccctgtct	ctactaaaag	tacagaatta	gccggcgtg	gttgcgcatt	cccataatcc	840
cagctactg	ggagactgtg	ggtagggaga	tcacttgaac	ctcagccgcg	gagggtgcag	900
tgagtcgaga	tcgcgcagg	gcactccagc	ctggcaaga	agagcgaac	tgggtctcaa	960
gttaaaaaaaaa	gaaagcaagg	aaagagtaat	ttacaacgaa	ggaaaaaaac	ccacagcaca	1020
cccttcgcgg	ctgtcagcgc	tctctgtatg	tcacagtggc	tgcgtgtct	tgggtgggt	1080
gaggtgtggg	gagcccgac	cctggccctg	cctccgcgc	cccgctcccc	ttctctct	1140
tactcggta	accatagcg	aggcccccgc	tcgtttcaga	tatgaatttg	ttttatagat	1200
tataaatatg	catataatgt	gtatgtataa	agcagaatgc	ctgcctttcc	tggttatitt	1260
ttgtaccata	ttgtaaatta	tattattat	tcttacca	tttggaaat	aaaaggttt	1320
ttggttat	aatataataa	gagctgttaa	actctgttt	aaatttccag	ttcaacttgt	1380
aaatgttttt	attgtgcata	aatacatact	aatgttgatc	taaaaaaaaaa	aaaaaaaaaa	1440
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1458

<210> 15

<211> 2005

<212> DNA

<213> Homo sapiens.

<400> 15

ggttgctggc	ccaggtgagc	gggcgcgcgt	gtccagggtga	gcggggcgcgt	ccccgcgcacg	60
gogctgcctg	cccgaggcgg	tacgttaaa	gacaggaga	tcctgagggc	cagccggaa	120
ggagggcgtgg	atatggagct	ggctgctgcc	aagtccgggg	cccgccgcgc	tgcctagcgc	180
gtcctggga	ctctgtgggg	acgcgcggcc	cgccgcggct	cggggacccg	tagagccgg	240
cgctgcgcgc	atggccctgc	tctcgccccc	cgcgctcacc	ctctctgtcc	tcctcatggc	300
cgctgttgc	aggtgccagg	agcaggccca	gaccaccgac	tggagagcca	ccctgaagac	360
catccggAAC	ggcggtcata	agatagacac	gtacctaacc	gccgccttgg	acccctgtgg	420
aggcgaggac	ggctctgtcc	agtataaaatg	cagtgaacgg	tctaaggcctt	tccctacgtt	480
tggttataaa	ccctcccccac	cgaatggatg	tggctctcca	ctgtttgggt	ktcatcttaa	540
cattggat	ccctccctga	caaagggttg	caaccaacac	gacagggtgt	atgaracctg	600
tggcaaaagg	aagaatgact	gtgtatggaa	attccat	tgctctcca	agatctgcgg	660
agatgtacag	aaaacactag	gactaactca	gcatgttcag	gcatgtgaaa	caacagtgg	720
gctctgttt	gacagtgtta	tacatttagg	ttgaaatcca	tatctggaca	gccaacgcgc	780
cgcatgcagg	tgtcattatg	aagaaaaaaac	tatctttaa	aggagatgcc	gacagctagt	840
gacagatgaa	gatggaaagaa	cataacctt	gacaaataac	taatgttttt	acaacataaa	900
actgtttat	ttttgtggaa	ggattatttt	gagaccttaa	aataattttat	atcttgcgtt	960
aaaaacctca	aagcaaaaaaa	agtgaggggag	atagtgggg	gagggcacgc	ttgtcttctc	1020
aggtatctt	cccaagcattt	ctcccttact	tagatgc	aatgtcttgc	ccaatatca	1080
aaacaagtgc	ttgttttagcg	gagaattttt	aaaagaggaa	tatataactc	aattttcaca	1140
accacattt	ccaaaaaaagg	agatcaaata	ttaaaatcat	cataatgtct	gttcaacatt	1200
atcttattt	gaaaaatgggg	aaatttacac	ttacaat	ttgtttacta	tggaaatttt	1260
aatacacatt	tatgcctaga	aggaacggac	ttttttttt	tatttttaatt	acacataata	1320
tgtattttttt	gtmcaacata	atatgttgc	tctctgttgc	ccgttggac	tatgtatgt	1380
tcacatttt	attaggacta	cttmcaagga	caaggttcc	attttccag	tgtaaaatt	1440
ggaaccatca	gctgataacc	tcgttagggag	caacccagg	atagctaagt	gttatgtat	1500
atgccttagaa	ggtgtatgtg	atgcgattca	gaagcatagc	cactccatt	ttatgagcta	1560
ctcacatgac	aaatgtcatc	ttttgtatata	acctttgcca	agtttagagaa	aagatggatt	1620
taatgagata	aatgaaaaga	tatattamct	aatatataaa	ggcactattt	gctgttatgc	1680

tttgttattt	atttcccagc	acttgttcct	tattgttagat	ttttaaaga	ctgtaaccctt	1740
ttactaactg	tggtcttact	aaaatttgc	cttgatactg	ctttcaaaa	agcccttaat	1800
tagagccaaa	aggatggaaa	aggcaagata	taaatgcctt	ttatagatct	cttatttaca	1860
ttgaaaatta	ttacccatatg	tttagagcaa	atccaacag	cttctgaaga	1920	
tgtctatgaa	tgttggaaaac	ttttcaatst	cttggatgc	tcakttaatt	cgcagacccg	1980
cttaacggat	taaacgcccc	cccccc				2005

<210> 16
<211> 943
<212> DNA
<213> Homo sapiens

<400>	16					
ggcacggact	ccgccccggcc	ccgaggggct	ctcccccggag	gctcagcccc	ctctgctccc	60
catggcaac	tgccaggcag	ggcacaaacct	gcacctgtgt	ctggcccccacc	acccacctt	120
ggtctgtgcc	actttgatcc	tgctgctct	ttggctctct	ggccctggcc	ttggcagctt	180
cctccatccc	cacaggactg	gcctcgccac	ccgtacatcc	cccaggactg	ggtctctttt	240
ttgagatctt	ttggccagct	gaccctgtgt	cccaggaaatg	ggacagtcac	agggaaagtgg	300
cgagggcttc	acgtcgtggg	cttgctgacc	accttgaact	tccggagacgg	tccagacagg	360
aacaagaccc	ggacattcca	ggccacagtc	ctggaaagtgc	agatgggatt	gaaaggatct	420
tctgcaggac	aactggctt	tatcacagcc	agggtgacca	cagaaggac	tgcaggaacc	480
tgccatatatt	tttagtgcgt	tccaggaatc	ctacccttca	gcoagccacc	catatcctgc	540
tcagaggagg	gggctggaaa	tgccacccotg	ageccctagaa	tgggtgagga	atgtgttagt	600
gtctggagcc	atgaaggcct	tgtgctgacc	aagctgctca	cctcggagga	gctggctctg	660
tgtggctcca	ggctgctgggt	cttgggctcc	ttctgtctt	tcttctgtgg	ccttctctgc	720
tgtgtactg	ctatgtgctt	ccaccccgccg	cgggagttccc	actggcttag	aacccggctc	780
tgagggcact	ggccctagttc	ccgacttgg	tctcaggtgt	gaatcaactt	cttggccctt	840
ggctctgagt	tgaaaaaggt	tttagaaaaa	gtgaagagct	ggaatgtggg	ggaaaataaaa	900
aagctttttt	gcccaaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaa		943

<210> 17
<211> 1503
<212> DNA
<213> Homo sapiens

<400>	17					
caggttcctc	tcagtamarc	ctcarsccga	ggttcccttc	ctcttgcattc	catgtgtgt	60
tttcaraggc	ggccatctt	ccctacttcc	agatccttgc	agggcagttg	gtggagggtt	120
ggaggcacc	cgtgtttgcc	tccatgaagc	cctgtgccag	tcactggct	gcaaggctga	180
ggaaaattgt	tccgtgtcag	aaagctccctc	agctcagagg	tgtgtgtacc	tcctgcgtgg	240
taggaaggca	ggggaaagag	gcccctgttc	tcctgttctc	tttgccttta	tgagacttga	300
gagtcgtgt	catctgtgcc	tttcgtgtct	ttttttcaga	ctccctgcga	caaggactgt	360
gtactgcatt	aatggggctg	agatagttga	tgttgcctgt	ggaatcctga	ttgagagccg	420
yaaacaggam	aaggcctgcg	agcagccggc	cctgggggg	gctgataacc	cagagcactc	480
ccctccctgc	tccgtgtcgc	ctcacacaag	ttctggagc	agcagtgagg	aagaggacag	540
tggaaacag	gcaactgrctc	caggcctcag	cccttccag	aggccgggg	gttccagctc	600
tgcctgtacg	aggagccctg	aggaggagga	ggaagaggat	gtgtgaaat	acgtccggga	660
gatcttttc	agctaggggca	taaactgtgc	actgaactgt	ctgcccagag	cagctggagg	720
acagctgagc	ttccactgtgt	gctgctggc	cgccccctcg	tgggaatggg	gctctctgtg	780
ctcctactt	tgtgccttct	tggccctggc	agattcacct	caggccagaa	gccccctggac	840
actccgggcc	ttggggctgc	cgttctgagt	gtgcggaaagg	caggactcaa	aatgagatcc	900
catttactc	cctctgtatg	tactgtgcc	tctcttggct	cttgaggctc	tggagtccca	960
attgtctgt	ttagtcagt	accaggttcc	agggaaaatg	atgtcatgt	gtggtccaa	1020
ttactggAAC	caaagagaca	gtactttgca	aagaaaagga	tcactgcccag	gtgcactgga	1080
attgtacag	tttagtccgc	atgatctctc	ctgaaggagg	aaggctgttt	caaaaatagt	1140
ttccatcatg	agtctatcaa	tgagctccca	cctctccagc	cagccatgaa	agcaaacgag	1200

ctgcccacag ttctctgcc	tgtctggag gttgaggcca	cagtgtata	actggtaagc	1260
cagacaggcc	tcctcccgca	agctgttacc	ttgctttcac	1320
agcttagctat	aaagcaagag	ggacaggagc	ccagaagaga	1380
accaggat	atgtctctgc	ctctgtttc	agtgtggctt	1440
aatca	atcagg	ttt	tggacaggaa	1500
ctcc	atacagg	ttt	tatatgaata	1503
			tga	

<210> 18
<211> 1512
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (207)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (209)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (521)
<223> n equals a,t,g, or c

<400> 18
gcagagcccc tgggtgtgag aagctcgct cccgtgggtt gcattggctc tgccctatct 60
ctgcctccag caccaggcc ggccgcagat ggcagtgtct ctggggacag cagctgcgaa 120
tgatgtccacg gggcaacgct gagctgctca ggctgaggcg gtgtgctcag cacagagccc 180
ccggaaactgg cttatgcagg cgctgancna aggccgcgc gatgcccac ttctggact 240
ggttcgtgcc ggtctacttg gtcatctcg teectatctt ggtgggcttc ggcgcctgca 300
tctactactt cgagccggc ctgcaggagg cgacacaatgt ggcgcacatgcg cgcgcgcctgg 360
tggaccgcse cctccgcaag acgctaattgg tgcgcgacaa cctggccttc ggcggccccc 420
aggtgtgagc cgacttgc aa agggatagg csgggccac cgggcgcctt ccccccagccc 480
gcccccccg cccagcccg agaccccaa ggcagaggga ngccggcttg ttggccctcc 540
acgctatccc tctgcagccct gggccctccc gacagaggcc ccagggtgcgc tgscagtgra 600
ggtgggac ttaggtgcct ggctggccca gggctgtctc tccgtgtcaa gccgactcac 660
ccagagccca cctcccaag ctcagggca tcctcgcgt gggcccgatg cctttgcrc 720
gcccgcact ctgcctccca ctggactcag gcatgtctat ggctgcctgt cctgaggcgc 780
cggagccctc atttctcgt gaagtccca gtcctccctgc ctccactaa tggcaccgjc 840
cctgcaactt taggcaggc gaagccaaacc caagaaaaga acactaaagac ctcgtttgga 900
gggatgtcag ctggggccag mcccagccca ccccgccggg ctcaggctt gaaactggtga 960
gggtgtgtgg tgggggtatg cagaggata agacctgtt agaggagagg gttgggtgagg 1020
agagagagag agagagagag agagagatc tggggggagc gggcaagcat ggggagatga 1080
gatgtgtata tggagagag agtgtgggg ccccaggcag ggcaggaggt ggtggaaacg 1140
gggtgaactc cgtgggctgt gtgaggactg tccatagtgg gtccmaaccc cctccctctg 1200
ctggagtttc cttagcccttc cccctcccyaa agactgwgc agcaggcagg agccctgtcc 1260
ctccctccct gtcctgtgcc acacttctgg ggccaaaccc agcccccctt agccaggccc 1320
tgccagactc caagccccc accctagaaccc ctcctgtt ggagactctg ttggcccaact 1380
ttggacacag attggcaacc tgcctcacm ckccccctw cgctgggct tccatcttaa 1440
tttattctca ataataaaaga cttcatgtg amaaaaaaaaaaaaaaa aaaaaaaaaaaa 1500
aaaaaaaaaa aa 1512

<210> 19

<211> 1655
<212> DNA
<213> Homo sapiens

<400> 19

ccacgcgtcc	ggcaaagaa	ttaaacctgg	tgtttggact	tcaacttagc	atggctagaa	60
ttgaaagtac	agtaaacatg	aacctcatgg	gatggctgt	ttctaagatt	gaagcttgg	120
tagttctgc	tggtcacaca	accctcgga	tcacacttat	gattgggggt	ataacgtgt	180
ttcttcaact	aatctgtgcc	ttggctcttg	cctacttgg	tcagagagca	gagagaatcc	240
ttcataaaaga	acaaggaaaa	acaggtaag	ttattaaatt	aactgtgt	aaggacttct	300
ccttacccct	gtggcttata	tttatcatct	gtgtctgcta	ttatgttgct	gtgttccctt	360
ttatggact	tggaaaagtt	ttcttacag	agaaaatttg	attttcttc	caggcagcaa	420
gtgcattaa	cagtgtgt	tatgtcatat	cagctccat	gtccccgggt	tttgggctcc	480
tggtgataa	aacaggaaag	aacatcat	gggttcttgc	cgatcatcag	ccactcttgt	540
gtcccacatg	atgctggcct	ttacgatgt	gaacccttgg	attgtatgt	gtcttctggg	600
actctctac	tcattgtctg	cctgtgcatt	gtggccaatg	gtggcatttg	tagtccctga	660
acatcagctg	gaaactgcat	atggcttcat	gcagttccatt	cagaatcttg	ggttggccat	720
catttccatc	attgctggta	tgatactgga	ttctcggggg	tatttggttt	tggaaagtgtt	780
cttcattgcc	tgtgtttctt	tgtcaacttt	atctgtggtc	ttactctatt	ggtgaatcgt	840
gcccagggtg	ggaacctaataa	ttattctgca	agacaaagga	agaaaataaaa	tttcccatac	900
tgaatgagaa	gtaaaaatga	atgtgtcaga	aatgggctt	aacacatcgt	tggtttggaaa	960
acttccat	taaaaattt	gatgttagtc	attaaaaaaa	ataatggact	ggaaagttaat	1020
atttatatcc	aaatatatactt	atttcaaagt	gtatttgta	ggcctgtttt	agectgtgtc	1080
tttttattt	tgtgttgcta	aagaatttca	cttttagtag	gctaatacac	aatgaaaagg	1140
tttagaaaattt	gctgtggAAC	atccaggtg	acttcaggaa	agacagtggaa	aaatggaaaa	1200
cgttggagct	tctgttgaga	taatcttcat	taggtatata	tcttagggat	acagcccttt	1260
ctttatctta	tagcaggaaa	aaaaaaactt	tgagggaaat	agaaggcgt	cgttacacaa	1320
aataaaacaat	ggcattgtca	taggccttcc	ttttactagt	agggcataat	gctagggaaat	1380
atgtgaagat	gttttttga	agtctcttc	tgatcacgaa	caatagcttg	cgctctactc	1440
tgttagttatg	tgtattgtccg	agcaatgacc	cttttcaatt	tcttattttct	gtgttactga	1500
ggaccctaat	cacttaggg	tgtatTTT	tagtataaac	tttctgtaca	gtttttctta	1560
tagtctaata	agtaaaaagt	gtccttcaa	ttatgataat	tgcctatgt	catggataaa	1620
ttaaaacact	gcacacggaa	aaaaaaaaaa	aaaaaa			1655

<210> 20
<211> 2525
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (5)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (10)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1354)
<223> n equals a,t,g, or c

<400> 20

tgacnctatn	gtaaggta	cctgcaggta	ccggccggga	attcccggt	cgaccacgc	60
gtccggctcg	ccaaacaagg	gttcatgaa	agtgttttc	tcttaaggt	aataaaaaaa	120

cagtggaatg gaaaaacagt gctgttagtca tcctgttaata tgctccttgt	caacaatgt	180
tacatctcg ctaggtgcctt tattcatgtc tttaagctca agtgcgcattct	tactagtgaa	240
gtattctgc aatgaaggta agttaaagact tggtatatgc atggagcact	tccatctaatt	300
cacacatctc tctcttgccctt ttggttctgt tatataaaatc atggaaataaa	taatgccttt	360
tgcttcattgt gaggatataaa gcatattaa attttgattat ttaaccttgc	attcctcaac	420
aagaaaaat gtttgataat ggatgaaatg tgagtcaatc agatcacaaa	atcaaaccct	480
tttgtgaaga accagtcgtt acatttgact gttattcaa tcaacaggtt	tttctggacc	540
tatagcaaaa ttgttaattt cgccttattt tgaagtagaa ggatataattt	ttttggtcac	600
ttggcatttg tgaggtactt actattgtaa ttattgtatc aatggtaagg	tgtcagcatt	660
atattgtgcg gtcataattgt atcaacagta taaattataa gcttgataa	gtatgtattt	720
aagaaatctt tttttatgtt gggatttaag caaacactt aattccacca	aactgtattt	780
agtacttctt acttagttt gaggtaaggg gtgggttgc cctccacatc	tgtgggtt	840
tctcgtttagg ttggaaacgaga gacttggaaa agaaaggac atagacaaaag	tatagagaaa	900
gaaaaaaagggg ggcccagggg accggcgctc agcacacggg	ggatctctgc	960
gttcmmttag tatttattgtc tcattattgg gtgtttctcg gagagtggg	tgtggcagga	1020
tcataggata gtagtgaga gagggtcaac aggttaaacac	gtgaacaaag	1080
cataracaak gtaaaagratt aagtgccttg cttttagata tgcatacaca	taaacatctc	1140
aatgcatttac aaagcagtat tgctgcccgc akgtcccacc	tccagcccta	1200
yccctatctc agtagatggc gcatacaatc gggttttata ccgagacatt	ccattgcccc	1260
gggacrrggca ggagacatg gccttcctct tgcataact gcaagaggcr	ttcttcctc	1320
ttttactata ctcctcagc acagaccctt tacnnggtgtt cgggctgggg	gacggtcagg	1380
tcttcctt cccacgaggc catatccag actatcacat ggggagaaac	cttggacaat	1440
acctggctt cctaggcaga ggtcsctcg gcytccrca	gttgggttgc	1500
cttgagatta gggagtgggt atgactctt asgagcatc	tgcccttcaag	1560
acaaaagcaca tcttgacccg cccttaatcc attaaaccct	gagtkgacac	1620
tcagagagca crgggttggg ggttaaggta	cagattgcag	1680
aatgtaaaca catcatgaaa atccattata ataactgcaa	taattgtatc	1740
gacacatatac atcatgatata gaccttaat ggtctactac	ttggagagtc	1800
ccaagtctct actgcaggta acatttacct	gccaggcaact	1860
gttcaaaagtg cataagatata agaccttgc	ctcaagactt	1920
aatgtaaaca catcatgaaa atccattata ataactgcaa	taattgtatc	1980
tgcagagatt gtctagggta	taaagttaattt	2040
tcaaataatggat cagagaacac	ttcgctgatt	2100
tgaacaaggg gattcaaaaac	agaagtcagt	2160
gcaggatatg ttgtgagaac	tgtactttt	2220
tagccaggca	ttgtgagaac	2280
tcacttgagg tcaggaattt	gagaccaggc	2340
gaaataaaaaaa aaattagctg	ttgttgcctt	2400
actgaagcag gagaatgaac	taatcccagc	2460
gcactccagc ctgggcaaca	tatgagggag	2520
gcccgc	tcatgccact	2525

<210> 21
<211> 1396
<212> DNA
<213> Homo sapiens

<400> 21

aagtctcgta tcgcgcgggg	gaggcgccgg	agccacgcgg	ctggcgccag	atccaggctc	60
ctggaaagAAC	catgtccggc	agctacttgtt	catgccaggc	acacactgtct	120
agctgctgtt	tgaattatct	gtgaatgtt	ggaagaggaa	tgcctcgatc	180
aattacccaa	ccaaagagaaa	tctgcaggat	ggactttctg	gtccctttct	240
ggcttcgggt	ctgtatgggtc	ttgttcttat	ctgcgtctgc	tcgaaaaccc	300
aggcctggcc	agggggaggag	cacagatatt	ttccctgtata	atagcttgc	360
agccrtgcatt	ggattgcattt	attacctttt	ccatacgaga	aaccacacct	420
gcacactggc	ttgcagggaa	ttggttatac	ttagtacacc	tgggaagttat	480
tcaggagctg	gagttgttctt	tgcattacct	tctctgcccc	tatctgtc	540

cctgtttttt ttacccctga cttgtggAAC caatccTGGC attataACAA aagcaaATGA	600
atttattttt ctcatgttt atgaatttGA tgaagtGATG tttccaaAGA acgtGAGGTG	660
ctctacttgt gatTTAAGGA aaccAGCTG atccaAGCAC tgcAGTGTGt gtaACTGGTG	720
tgtgaccCGT ttGACCACt acTGTGTTG ggtGAACAAC tgcATCgggg CCTGGAACAT	780
caggtaCTTC ctCATCTACG tcttgaccCTG gacGGCCtG gtcGCCACCG tcGCCATTGt	840
gagcAccACT tttctGGTCC acttGGTGTGt gatGTCAGAT ttatACCAGG agacttACAT	900
cgatgacCTT ggACACCTTc atgttatGGA cacGGCTTT cttaTTCAgT acCTGTTCT	960
gactttCCA cgGATTGtCT tcatGCTGGG ctttGCTGTG gtTCTGAGCT tcCTCCTGGG	1020
tggctacCTG ttGTTTGTCC tGTatCTGGC ggCCACCAAC cAGACTACTA acGAGTGGTA	1080
cagagGTGAC tggGCGTGTG gCcAGCtGtG tcccCTGTG gcCTGGCtC CGTCAGCAGA	1140
gcccccaAGTC caccGGAAcA ttCACTCCCA tggGCTCAG agCAACCTTC aAGAGATCT	1200
tctacttGcc ttTCCATGTC atGAGAGGAA gaaACAAGAA tgacaAGTGT atGACTGCCT	1260
ttgagCTGTA gtTCCCgtt atttACACAT gtggatCCTC gtTTCCAAA aaaaaaaaaa	1320
aaaaaaaaaaaa aaaaaaaaaa aaaactcGAG gggggggcccG gtacCCAAAtt CGCCCTGGAG	1380
ttcaagtAGA catCAA	1396

<210> 22
<211> 1069
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (508)
<223> n equals a,t,g, or c

<400> 22	
ggcacGAGCA cAGCCTCAGG CCCTGCCCCA GACCTGcAGA ATCAGAAACT CTGGGGTAG	60
gcCTGTTAT CTGCTGTAAc AGACCTTCCA GTGGGTTCTG ATGCCCTCTA GAGCAGGAGA	120
accACTAGCT TAGAGGTGc AGATGTTG GCATCTGc ATTGTGTTA GTTCAGAGGA	180
atGGCTGACC CccATGTCTC ATTCTAAGC ttCAGGcAGC TTTCTCTG ggcAGCTGTC	240
attCTGTG ACCCTCTTCA tCTCCCCGTG gaggCTCCAT GCCTAGAGGT ggtCTTCAA	300
cagaAGAATG GCAAAGATAA ttGTCCTGTG ttttACCTG ACCCCATTcC TTAAAGAGGG	360
tcacttCTTg GCCATTATC TAaaaACCAA tGTcATAGTT CTGTGATTcC ACTATCAGAC	420
agtGccACGT CCAAGGCGCG ggCTTNAc CTCCCTGGAA gagAGACTGT gCTGTCTGT	480
cttCCCTGTGt tCTCCAGTCC cacGCTCCCA CGGACCCACG CCCTTGGAGA CTCCCTCGGT	540
gtcccAGGGC ttCTGGTGTG tTCAGAGACC tCCACACTCA ACgACCACTG GTGCTGcAGA	600
agggCCGGTG CCTACATCC AATTAACAGA CGCTTTCCC ATCTAATGCC tCTTGCCTTC	660
tcctaACACC ACCTGGGAG tGTTTATGt tATTCTAAGT GAATTCTACT GTGTGAAAAAA	720
attCACACCT GTGTCCCGAG CGATTTGGGA GGCGGGGCG GGTGTATCAT ttGAGCCAG	780
gagTTGAGG CTAGCTGGG CAGGATGGTg AAACCCCTGtC tCTATAAAGA AATTtTAA	840
atAGCTGGG CATAGTGGCA CGTGCCTGTA GTTCCATCTA CTGGGGAGGC TGGGGTGGGA	900
ggatGcATG AGCCCGGGAG tttGAGGCTG CAGTGAAGTg TGATGcAGC ACTGCACTCC	960
agtCTGGGCA ACAGAGCAAG ACCCTGTCTC TTTAAACAAA AAAA	1020
	1069

<210> 23
<211> 1658
<212> DNA
<213> Homo sapiens

<400> 23	
ggcacGAGCC GGCCTGCCAG AGCCATGCC CTGACTCCTC AGCTTCAAA TCAGGGGTCT	60
caggACAGAG GATGCTGGGT GGGCTCAGAG CTCACTAGGG GGGCTGTG TGAGAGGGGA	120
tgcCTTCTGG ATGCCCTCAT CCTCCCTGGG GCTGGGGTCT CCCTCAAGGC CACCCAGCTC	180
cttCCCTTGTt TTGCTGCTGt tactCCTGCC GCCTGCTGCC TTGCCCTGC TGCTCTTCT	240

cttgacttc ttccctccca	gggcagctgt	gtctcccttc	ttgcggacc	actgctctgc	300
caggcaacct	agggtgtgga	ggagagagac	cctcaacaga	agtgcctcag	360
ctggcaagg	agcaactgagc	agggagccgt	gggagtagca	actgggactg	420
cagccctccc	gcctcctgcc	tctcgctgtg	gccaccaggc	ccctctgggg	480
tgtctcccc	tcattctgca	cctggttca	gtgactctta	cttcacccatg	540
caaggcttca	aagggcagag	atccataaat	gccctctggt	tctccctcta	600
cccacccccc	ccaagggaga	gcaaacaat	catccagagc	cagectgccc	660
aaactcaactg	gtgtctttc	ccttcagcac	gtggttggcg	cagaggaatt	720
ttcctccctgg	tggatccctgt	accagtcctc	gatcacctcc	tcaaacttcc	780
tcacacttgt	taatcataat	acccatatt	ggcaaagccc	cagcacctga	840
gaggagctca	gcttaagcct	cgcaaaaaac	tgcaaggtag	cagtggcacf	900
gaaaactgag	gcccagaggg	tgaatgacc	tgacccaaagg	caccccccggcc	960
tcctcagagc	agaccccaatc	cccaccagcc	cttcaactggg	cacagcaacc	1020
ctgaaggggcc	tgtacactgtct	tcttgagggt	agccacttct	gcagaagttct	1080
ctcatagggg	atgtccatca	ccacccatgc	ccctttgtgt	accagggtgt	1140
aaaggctct	gacatccct	ggaagaagcg	accagatatg	gcaggccggag	1200
tccctccac	cctcgctctcc	cagtgggtgc	taagaaccca	gtataagac	1260
cgcctctaa	ggatccctcat	cctttttttt	ttgagaagga	gtctcaactt	1320
ttggagctc	tcagctca	gcaacccttg	cctccagggt	tcaagcgatt	1380
agcctccaa	gcaagctggg	ctacaaaggc	gtgcaccat	acccggctaa	1440
gttgggttt	tgtcatgtt	gtcaggctgg	tctcgaaactc	ctagcatcaa	1500
cacccatggc	tcccaaagtgc	ctgagattac	aggcgatggc	caccgcaccc	1560
ttgacccatgc	cttcctcttc	cctcttttag	gcctgttcc	cacaacccct	1620
cccctgatct	gcctctgac	acccatcatgc	ttcaaaaa	gcacatataac	1658

<210> 24
<211> 1077
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1036)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1038)
<223> n equals a,t,g, or c

<400> 24

ggcacgaggg	gaaagccatg	ctcccaggac	tccttccttg	cagccctaaa	tcggctctgt	60
cggaaattc	cgcgccttag	aaacccacgc	ttgggtgtaa	cttattattg	ttcttcctga	120
cctacttcct	gtttatca	tccgggttca	tcattttggc	atttcggta	tcgggttgaa	180
actattgaag	cccgcttca	gtttcttttc	cccatttcc	cttggaaagg	aagacttctg	240
gcttcctta	aatctccgtt	ctctgggtaa	ggggagtcga	agccctgtc	atgaggaacg	300
gaaatgcgag	ggcctcggtt	gttactctaa	aatcccccct	cagttgcac	gccggaaact	360
gcgatttcctg	cagcggaaaga	ggcgtgatct	ggccctcgac	tcgtatgtc	cactaacaat	420
atgtcgacc	cacggaggcc	gaacaaaggc	ctgaggtac	agccccccgc	gagcgaatgt	480
aacccggcc	tggacgaccc	gacccggac	tacatgaacc	tgtggccat	gatcttcagc	540
atgtcgacc	tcatgtttaa	gctgaagtgg	tgtgcttggg	tcgctgtcta	ctgctccctc	600
atcagctttg	ccaaactctcg	gagctcgag	gacacgaagc	aatgtatgag	tagcttcatg	660
ctgtccatct	ctgcccgttgt	gatgtccat	ctgcagaatc	ctcagcccat	gacgccccca	720
tggtgatacc	agcctagaag	ggtcacattt	tggaccctgt	ctatccacta	ggcctgggt	780
ttggctgcta	aacctgtc	tttcagctc	catccctggac	ttccctgaat	gaggccgtt	840
cggtgcccc	agctggatag	agggAACCTG	gcccttcct	aggAACCCAC	ctaggcttac	900
ccctctgccc	tccctcccc	tgcctgctc	tggggagat	gctgtccat	tttcttaggg	960

tattcatttgc	ctttctcg	gaaacctgtt	gttaataaag	tttttactc	tgaaaaaaaaaa	1020
aaaaaaaaaa	aaaaancncg	agggggggcc	cggAACCCAA	ttcsccggat	agttagt	1077

<210> 25
<211> 1205
<212> DNA
<213> Homo sapiens

<400> 25	ccccacgcgtc	cgcagaggca	ggcaaatagt	ggagttctgg	ctggccaag	cagcctagaa	60
	ctcaaagtcc	atggccctt	ctgggcctgg	agaaatttgg	tgttatagc	accaggcagc	120
	ccttgggggt	gggggacagc	aataggagga	cctctttt	ctctacactc	tcctttggct	180
	cccgagatc	tggcaggccc	tggctgggg	cataagatta	gatgagggtt	agctgttga	240
	gaatgaagct	gtgttgggg	aagaattag	gttgtaccgg	aagatcaacg	aggttgtct	300
	gtcaggaaat	gagggtggat	ttggggccaa	gtgaggctgc	attattagat	aatgagggtt	360
	gtactgtcag	gggatgaagt	gtacttgtag	tagagatgac	gtccctgtgg	atcagtcggc	420
	ttttgtcca	tcagagaaca	cagccacacc	acaggaggaa	ggagagtgtc	cgactcagag	480
	gataaaatgag	gtgttccctgc	tggataatg	agggggcccg	tcaggtgaat	ggagtgtctgt	540
	tagcaaatg	gtttgtactt	gctggataaa	tggactgtt	gtgctggata	aatgggggtt	600
	tgctgtcagg	tgaatgcatt	actgctcg	ggtgaaggc	atccctggaa	taatgagggt	660
	gtcctgtctt	atagatgagc	tgccaccacc	aatggatca	gaccctgtcc	atgaaggagg	720
	caccatcagc	aacgcacgagg	tatcctgtt	cccactgggg	ctctggagc	gtcttctggc	780
	ccaggggaaa	ctcgggtgtt	gccaccctgg	gttattcaag	tctctgtgg	gagcagggtt	840
	gggggctggg	gagggcaggc	agtcgcattt	tgcaccgtgg	gacccttcct	tcacccccc	900
	tggatgcctt	actcctctcc	ctggcacccc	tcagtggtt	agactgttcc	ggacatttcc	960
	accccaactgc	ctgcttctca	tcctgcctgt	gtcttcttcc	tgcccagttt	ggaaaagccc	1020
	ctattatgtt	tcagccactc	tgcccagttt	tatataatct	ccctataaca	cagtattact	1080
	cctcttgca	catacacact	ttcttttatt	cattcatcca	ttcatttatt	tgacaaacat	1140
	ttaagtgtct	agtagtgc	aaacacatga	gttacagttt	taaaaagaat	aaaaaaaaaa	1200
	aaaaaa						1205

<210> 26
<211> 1674
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1663)
<223> n equals a,t,g, or c

<400> 26	ccccgaggc	tgagtccctt	ccctgtcttt	cactctctg	gcatcggtgg	ttttacttct	60
	tcgattgaac	cctgttttcc	cgaccctt	gggaggccgc	cttcttcagg	cgccctccctt	120
	ctctccacga	gctcgctctg	acagctgagg	aactggcaag	atctgtctac	ccagagggtt	180
	aatgggtatc	tttccggaa	taatccta	ttttctaagg	gtgaagttt	caacggcggc	240
	cgtgattgtt	agcggagtaa	gcaaacactt	ccattgtatt	agtcaccaga	aaagtaccac	300
	tgttaagtcat	gagatgtctg	gtctgaat	gaaaccctt	gtatatggcg	gccttgcctc	360
	tatcggtct	gagtttggg	cttccctgt	ggaccttacc	aaaacacgc	ttcaggttca	420
	aggccaaagg	attgtatgccc	gtttcaaaaga	gataaaatat	agagggatgt	tccatgcgt	480
	gtttcgatc	tgtaaagagg	aagggttatt	ggctcttat	tccatgcgt	tcctgcgtt	540
	gctaagacaa	gcatcatatg	gcaccattaa	aattgggatt	taccaagct	tgaagcgctt	600
	attcgtagaa	cgtttagaag	atgaaactct	tttaattat	atgatctgt	gggttagtgc	660
	aggagtgtat	tcttccacta	tagccatcc	caccatgtt	ctaaagatc	aatgcagggc	720
	tcaagaaagg	ttgttccaag	ggagcatgt	tggaaagctt	atcgatatat	accaacaaga	780
	aggcaccagg	ggtctgttgg	gggggtgtt	tccaaactgt	cagcgtgt	ccatcggtt	840

aggagttagag	ctaccagtct	atgatattac	taagaagcat	ttaatattgt	caggaatgat	900
ggcgataca	attttaactc	acttcgttcc	cagcttaca	tgtggtttg	ctggggctct	960
gccctccaac	ccgggtgatg	tgttgcAAC	tcgcatgatg	aaccagaggg	caatcgtggg	1020
acatgtggat	ctctataagg	gcactgttga	tggtatttt	aagatgtgga	aacatgaggg	1080
cttttttgcA	ctctataaaAG	gattttggcc	aaactggctt	cggcttggac	ccttggAACat	1140
cattttttt	attacatacg	agcagctaaa	gaggcttcaa	atctaagaac	tgaatttat	1200
gtgagccccag	ccctgcocagc	ctttctactc	cttgcocctt	ttcccggtt	ctaatgtatt	1260
ttgacaatgt	tgtaagtgtt	taccaagccg	ttggtctct	aaggggcttc	tgatggaaaga	1320
acagtgggggt	ggttccaaagt	tattttctatg	tttggtgttac	catgttaact	tttcccccggag	1380
agaaaagtgtt	aacatttgaga	ctctggcccc	agattggtat	cttctatgaa	gatggataact	1440
gatgggtgac	attgaaaacg	gcctgcgttcc	caaatgtgtt	ttaatgtata	tggtttagccc	1500
cagacttggg	ctagagcaga	aggcataggc	cagggtgggt	attgtctatat	gtgttacaga	1560
cctcggttcc	cattaaagta	tttattggca	gaataaaaaa	aaaaaaaaaa	aaaaaaaaaa	1620
aaactcgagg	ggggggcccg	tacccaaattc	gccctatgtt	gantcgaatg	ggct	1674

<210> 27
<211> 1965
<212> DNA
<213> *Homo sapiens*

<220>
<221> SITE
<222> (333)
<223> n equals a.t.q. or c

<220>
<221> SITE
<222> (1961)
<223> n equals a.t.g. or c

<400> 27
ggatccctcgc ggcggcgccg gtgtttacag cctgagaaga gcgttcgc cgggagccgg
ggcgccatc gagaccacc caaggcgcgt cccccctcgc ctccccagcgc tcccaagccg
cagcggccgc gcccccttcag cttagtcgcgt cgctcgctt gttccctcgc tgccggctgc
gcatggcggtt ggcgttggcg gcgcgtggcg cggtcgagcc ggctcgccg acccggtacc
agcgttgca gaatgaagaa gagtcgtggag aacctgaaca ggctgcaggat gatgttcctc
cacccttacag cagcatttcgc agagagagcgc cancatatt tgactacaag gatgagtctg
ggtttccaaa gccccccatct tacaaatgttag ctacaacact gcccaggat tatgtaaagccg
agaggccaa ggctgaagact atatatccctt tggttcctgg gagagatgag gattttgtgg
gtcgggatga ttttgcgtat gctgaccagc tgaggatagg aaatgtatgg attttcatgt
taactttttt catggcattc ctctttaact ggattgggtt ttccctgtt tttgcctgaa
ccacttcagc tgcaggaaagg tatggggcca tttcaggatt tggctctctt ctaattaaat
ggatccgtat tgtcaggattt tcacccattt tccctggata ttttgcgtt cagtaactggc
tctgggtgggt gtcccttgc tttaggcattt tcctgtttt cagaggattt atcaattatg
caaaaagttcg gaagatgcca gaaactttct caaatctcc caggaccaga gttctcttta
tttattaaag atgttttctg gcaaaggcct tcctgcattt atgaattctc tctcaagaag
caagagaaca cctgcaggaa gtgaatcaag atgcagaaca cagaggaata atcacctgt
ttaaaaaaat aaagtactgt tgaaaagatc atttctctt atttgcgtt agtgtaaaa
ttttaatagt taatgcagaa ttctgtatc attgaatcat tagtgcgttta tggttgaaaa
agctcttgca atcaagtcgt tgatgttata ataatgcctt atatattgtt tgtagtcatt
ttaagtagca tgaggccatgt ccctgtatgc ggtaggggc agtctgtt tatttcattcct
ccatctcaaa atgaacttgg aaaaatggat tgaatgtat gtataatgtt gggcatttt
aagggtttt ctcaaaaatgg aacttttgtt tgatgttgc ttttgcaca taatccatatt
ttgctgttca agttaatcta gaaatttttattt caattctgt tgaacacctg gaagcaaaaat
catatgtccaa aaatacattt aagggtgtggt caaaaataag tctttatttg gtaaataata
agcattaatt ttttatagcc ttttgcgttca attctgcggt accttattgt acttaaggga
ttctaaaggt gttgtactg tataaaaacag aaagcactag gatacaaaatg aagctttaatt

14

actaaaatgt	aattcttgcac	actctttcta	taatttagcgt	tcttcaccccc	caccccccacc	1620
cccacccccc	ttattttcct	tttgtctcct	ggtgattagg	ccaaagtctg	ggagtaagga	1680
gaggatttagg	tacttaggag	caaagaaaaa	agtagcttgg	aactttttag	atgatcccta	1740
acatactgt	ctacttgcct	ttacaatgtg	tttagcagaaaa	ccagtggtt	ataatgtaga	1800
atgtatgt	ttctggccaa	gtggtaattc	atcttggtt	gctatgttaa	aactgtaaat	1860
acaacagaaac	attaataaat	atcttgcgt	tagcacctt	aaaaaaaaaa	aaaaaaaaaa	1920
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	naaaa		1965

```
<210> 28  
<211> 1863  
<212> DNA  
<213> Homo sapiens
```

<400>	28	60					
gactaggccg	cgagcttagt	cctggggagcc	gcctccgtcg	ccggccgtca	agccgc	cccta	120
ttagattatc	ttaacaagaa	aaccaactgg	aaaaaaaaat	gaaattcctt	atttcgc	cat	180
tttcgggtgg	tgttcacctt	ttatcccgt	gctctggaa	agctatatgc	aagaatggca		240
tctctaagag	gactttgaa	gaaataaaag	aagaaatagc	cagctgtgga	gatgttgct	a	300
aagcaatcat	caaccatgt	gtttatggta	aagcccagaa	cagatcctat	gagcgattgg		360
cacttcttgtt	tgatactgtt	ggacccagac	tgagtggctc	caagaacta	aaaaaagcca		420
tccaaaattat	gtacccaaaac	ctgcagcaag	atgggctgga	gaaagttcac	ctggagccag		480
tgagaatacc	ccactgggg	aggggagaag	aatcagctgt	gatgtggag	ccaaagaattc		540
ataagatagc	catecctgggt	cttggcagca	gcattggac	tcctccagaa	ggcattacag		600
cagaagttct	ggtggtgacc	tcttcgatg	aactgcagag	aagggcctca	gaagcaagag		660
ggaagattgt	tgtttataac	caaccttaca	tcaactactc	aaggacggtg	caataccgaa		720
cgcagggggc	ggtggaaagct	gc当地gg	gggctttggc	atctctcatt	cgatccgtgg		780
cctccttc	catctacagt	cctcacacag	gtattcagga	ataccaggat	ggcgtgcccc		840
agattccaaac	agcctgtatt	acgggtggaa	atgcagaaat	gatgtcaaga	atggcttctc		900
atgggatcaa	aattgtcatt	cagctaaaga	tgggggcaaa	gacctacca	gatactgatt		960
cctcaacac	tgttagcagag	atcaactggg	gcaaataatcc	agaaacagg	ttgtactgtca		1020
gtggacatctt	ggacactgtt	gatgttgggc	agggtgc	catggatgtggc	gggtggagcc		1080
ttatatcatg	ggaagactc	tcacttatta	aatagatcttgg	gctgcgtcca	aaaggagactc		1140
tgccgttgtt	gctctgtact	gc当地aaac	aaagggtggat	tgggccttc	cagttattatc		1200
agttacacaa	gttaaatatt	tccaaactaca	gtctgtgtat	ggatgtctac	gc当地aaact		1260
tcttacccac	tgggctgca	ttcaactggca	gtgaaaaggc	caaggccatc	atggaggagg		1320
ttatgagcct	gctgcagccc	ctcaatata	ctcagg	tcctt	gagccatgga	gaagggacag	1380
acatcaactt	ttggatccaa	gctggaggtgc	ctggagccag	tctacttgc	gacttata	aca	1440
agtatttctt	cttccatcac	tcccacggag	acaccatgac	tgtcatggat	ccaaagcaga		1500
tgaatgttgc	tgctgtgtt	tgggctgtt	tttcttatgt	tgttgagac	atgaaagaaa		1560
tgctgcctag	gtccctagaaa	cagtaagaaa	gaaacgtttt	catgttctg	gccaggaaatc		1620
ctgggtctgc	aacttggaa	aactctctt	cacataacaa	tttcatccaa	ttcatcttca		1680
aagcacaact	ctatttcatg	ctttctgtt	ttatcttttct	tgatacttcc	caaattctct		1740
gattcttagaa	aaaggaat	tttccccc	cctcccccacca	catagaatca	acatatggta		1800
gggatcacag	tggggcatt	tctttatatac	acctcttaaa	aacattgttt	ccactttaaa		1863
agtaaacact	taataaattt	ttgaaagatc	tctgaaaaaa	aaaaaaaaaa	aaagggcggc		
ccgc							

<210> 29
<211> 1626
<212> DNA
<213> *Homo sapiens*

```
<400> 29
cccacgcgtc cggagccggg agccggtcgc gggggctccg ggctgtggga ccgctggcc 60
cccagcgatg gcgaccctgt ggggaggcct tttccgttgc ggctccttgc tcacgcgtc 120
gtgcctggcg ctttcgttgc tgctgtggc gcactgtcaq acqccgc当地 gtattgtc 180
```

tcatgttgtg gagccccatgc ctgtcgaaaa	gcctgtatgtaa	gaagcataact	gtctacgcgt	240	
tgaatcaaaa tatgaagaaa	gaagctctgt	cacaatcaag	tttaccattat	300	
ctccatTTTg ggccttctac	ttctgtatcat	ggtatatctt	actctgggtt	360	
gaaggaggcgc ctctttggac	atgcacaggat	gatacagagt	gtatgtata	420	
ccagcctttt gcaaatgcac	acgtatgtgt	agcccgcgt	cgcagtcgag	480	
gaacaaggta gaatatgcac	agcagcgctg	gaagcttcaa	gtccaagagc	540	
tgtctttgac cggcatgtt	tcctcagcta	attggaaattt	gaattcaagg	tgactagaaa	600
gaaacaggca gacaactgga	aagaactgac	tgggtttgc	tgggttcat	ttaataacct	660
tgttgatttc accaactgtt	gctggaaattt	caaaaactgga	agcaaaaaact	tgcttgattt	720
tttttttttg ttaacgtaat	aatatagagaca	ttttttaaaag	cacacagtc	aaatcagcca	780
ataatctttt ctttattgtga	cttttactaa	taaaaataaaa	tctgcctgt	aattatcttg	840
aagtccctta ctttgcacaa	gcaactcttt	tttcacccaca	tagttttaac	ttgactttca	900
agataatTTTt cagggtttt	gttggatgtt	ttttttgttt	gttggatgtt	gtgggagagg	960
ggaggatgc ctgggaagt	gttacaact	tttttcaagt	cacttacta	aacaaaactt	1020
tgtaaataga ctttaccttc	tattttgcag	tttcattttat	attttgcagt	gttagccagcc	1080
tcatcaaaga gctgacttac	tcatttgcact	tttgcaactga	ctgtattatc	tgggtatctg	1140
ctgtgtctgc acttcatggt	aaacgggatc	taaaatgcct	ggtggctttt	cacaaaaaagc	1200
agattttctt catgtactgt	gatgtctgt	gcaatgcac	ctagaacaaa	ctggccattt	1260
gctagttac tcataaagact	aaacatagtc	ttgggtgtgt	tgggttact	catcttctag	1320
tacctttaag gacaaatcc	aaggacttgg	acacttgca	taaagaaaattt	ttatTTTaaa	1380
cccaaggcctc cctggattga	taatataatac	acatttgcac	gcatttccgg	tcgtggatgg	1440
aggcagctgt ttgagctcca	atgtgtgcag	ctttgaacta	gggctgggtt	tgtgggtgcc	1500
tcttcgaaa ggtctaacc	ttatggata	actggctttt	ttcttcctct	ttggaaatgt	1560
acaataaaaaa taatTTTga	aacatcaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1620
aaaaaaaa					1626

<210> 30
<211> 605
<212> DNA
<213> Homo sapiens

<400> 30					
ccacgcgtcc gcccacgcgt	ccggggaaatg	accttggaga	ttttagcaga	gagtggcat	60
gaggagcggc ctgctggcca	ggggccggat	gagcccaaca	tgaaccctaa	gcttggggac	120
ccaaaggcgc cccacaccc	cttccctgtgg	tttacctccc	catacaagac	catgaagt	180
atccctgtgc ggcgtttccg	gtggggccatc	atcccttca	tcatcccttt	catccctgt	240
ctgttcttgc ccacatcttcat	ctacgccttc	ccgaactatg	ctgcacatgaa	gctgggtgaag	300
cccttcagctt gaggacttgc	ctggccctgt	gaagggggccg	tgggggtcccc	tccagcatgg	360
gactggccgt cctccctccgc	ccagctcgcc	gagctccctcc	agaccccttca	ggcctgtatt	420
tcctggccagg gtgggcagac	agacagatgg	accggggccac	actcccttgc	ttgctaaat	480
ggagctctga gatcacccca	cttccatcat	ttccctctcc	cccaacccca	cgcttttttgc	540
gatcaactca gacatatttc	agtataaaac	agttggaaacc	acaaaaaaa	aaaaaaaaaa	600
aaaaaa					605

<210> 31
<211> 931
<212> DNA
<213> Homo sapiens

<400> 31					
gagagtgcct aagggggggt	gaaagaggac	gtgttaccca	ctgccccatgca	ccaggactgg	60
ctgtgttaacc ttgggtggcc	cctgctgtct	ctctgggctg	cagagtctgc	cccacatgt	120
gccatggcct ctgcaactgc	tcagctctgg	tccaggccct	gtggcaggac	acacatgggt	180
agcctagccc tggacatca	ggagactggg	ctctggctct	gttccggccctt	tgggtgtgt	240
gtggatttctc cctggccctc	agtgtgccca	tctgtaaagg	ggcagctgac	agtttgtg	300
atcttgccaa gggccctgt	gtgtgtgtat	gtgtgtgt	gttccatgtgc		360

gtccatattt aacatgtaaa aatgtccscc crckcgtccg cccaaacatg ttgtacattt	420
caccatggcc ccctcatcat agcaataaca ttecccactgc caggggttct tgagccagcc	480
aggccctgcc agtggggaaag gaggccaagc agtgcctgcc tatgaaaattt caacttttcc	540
tttcatacgt ctttattacc caagtcttcc cccgtccatt ccagtcaa at ctgggctcac	600
tcacccccgc gagctctcaa atccctctcc aactgcctaa gggcctttgt gtaaggtgtc	660
ttaatactgt cttttttttt ttttaacag tgttttgttag attcagatg actatgcaga	720
ggcctgggg acccctggct ctggggccgg cctggggctc cggaaattcca aggcccagac	780
ttgcgggggg tgggggggta tccagaattt gttgtaaata ctttgcatat tgtctgat	840
aacacaaaca gacctcaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa	900
aaaaaaaaaaa aaaaaaaaaaaa aaaggcgcc c	931

<210> 32
<211> 1407
<212> DNA
<213> Homo sapiens

<400> 32	
gaattcggca cgaggcgagg ctcagaagac gatgcggggc tttgtgtccgg ctttcttgc	60
gcatgtactc ccgtctcgagg gagcttgcgt caccaggcgcc gcccaggatct tcacagcaca	120
actgcctggaa aggccagggtt cgagaaaggag aggccggatgg catggcaggc aagggggacc	180
gatgtgtgc agctcacacc actccagaac ctgacaaggc accaggcagg cccttgc	240
ggagcatgtc tttgtgtccgt cctgcacatt ccagaagccc tcatggaaag	300
ggatcgagcc aggcagactc ctgcaggatg gggcaggtag tttattcaaa gagaactctg	360
tatcccatag gcccaggctc tcccttcgtc tggcggtggc tttgtgtccgc cagtggtgtc	420
tcctggctca gcagaaacat ccatttgagt tggcatccct gttagggatcc cagagcggt	480
taaggcttct tttgtgttggt agggatggct gtgggggtgc tttccaggagg gggccaccat	540
tgccgcattctt acttctagac tcccaaaggaa gcccaggctc aggccaggctt gggccagagt	600
cacgcgtggca accacaggatg tggaaaggcag tcgtatttc tctctctctc tctctctctc	660
agtatccatg agaggatgtaa aacatattttt ctcttataa atgtctttt acaaattatg	720
tgattatctg gaagctctaa gatggagagca aatgcctgtat cactgtggcc aatgtcaga	780
tactaaaaggcc catttttggc cgggcatgtt ggctccggcc kgtatccca gcactttggg	840
aaggccaaatg gggtaatca cctgaggatc ggagttcaag accagctga ccaacatggg	900
gataccccgt ctctactaaa aatacaaggcc gggcggtggc gccatgcct gtaatccca	960
ctactcagga ggctgaggca gggaaatcac ttgaactcgg gaggcagagg ttgcagttag	1020
ctgagatcgc gccattgcac tccaggctgg gtgacagagc aagactctgt ctcataaaata	1080
aatacaaaggcc ccatttttcc agagtcttgcgtt gccttaaaaata aaacacaccc tctgtgt	1140
ggaagactgt gcaatggcac agccgcaggag cttgggtttgg gagggttgaag tgctctgggg	1200
agaatctcgta gatcatccctc agaaaaaggctc tgccctggtg ttctaccaga aaaacgtctc	1260
ccaaatcccc agggaaaggctg tccacagtag tcccccctta tccacgggtgt cactttccat	1320
gggtttagt atctgcggc aaccacggc tgacaatattt aaatggaaaa ttcttcaaac	1380
agttaaaaaaaa aaaaaaaaaaaa aactcga	1407

<210> 33
<211> 1526
<212> DNA
<213> Homo sapiens

<400> 33	
ggcacgagaa aaaaccttca ggcggcccat ggttatgccc aagaggacag agaacgaatg	60
cacagaaaaata ttgtcagccct tgcacagaaat ctccctgaact ttatgattgg ctctatcttgc	120
gatttatggc aatgttttctt ctgggtttac attgggttctt cattgaatgg tactcgggaa	180
aaaagagttc cagcgcaactt ttccaaacaca tcactgcatt atttgaatgc agcatggcag	240
ctattatcac ctacttgcgt agtgatccag ttgggtttctt ttatattcgt tcatgtcgag	300
tattgtatgtc ttctgactgg tacacgtatc ttccaaaccc aagtccagat tacgttacca	360
cagtacactg tactcatgaa gccgtctacc cactatatac catgttattt atcttattacg	420
cattctgctt ggttataatg atgctgcttcc gaccccttctt ggttgaagaag attgcgtgt	480

ggtagggaa atctgatcga	tttaaaaagta	tttatgtgc	acttacttc	ttcccaattt	540
taaccgtgc tcaggcagt	ggtgaggccc	ttttatatta	cgcctccca	tacattatat	600
tagtggatac ttgggtact	ctggctgtgt	acatgtctgc	ttctgaaata	gagaactgt	660
atgatttctt ggtcagaaag	aaaagactta	ttgttcttctt	cagccactgg	ttacttcatg	720
cctatggaat aatctccatt	tccagagtttgg	ataaaacttga	gcaagatttgc	ccccctttgg	780
cttttgttacc tacaccagcc	ctttttact	tgttcactgc	aaaatttacc	gaaccttcaa	840
ggataactctc agaaggagcc	aatggacact	gagtgttagac	atgtgaaatg	ccaaaaaacct	900
gagaagtgtc cctaataaaa	aagtaaatca	atcttaacag	tgtatgagaa	ctattctatc	960
atatatggga acaagattgt	cagtataatct	taatgtttgg	gtttgtctt	gtttgttta	1020
tggtagact tacagacttg	gaaaatgca	aactctgtaa	tactctgtta	cacagggtaa	1080
tattatctgc tacactggaa	ggccgcgtt	aggcccttgc	ttctctcaac	agttcagctg	1140
ttctttaggg caaaatcatg	tttctgtgt	cctagaaatg	tgttcccatt	ttattaagaa	1200
aagctttaac acgtgtatc	tgcaagtccctt	aacagtggcg	taattgtacg	tacctgtt	1260
gtttcagttt gtttttacc	tataatgtat	tgtaaaaaaca	aaacataacttg	tggggctga	1320
tagcaacat agaaaatgt	tatattgtt	tttgcattct	atttatttcc	atcaatacag	1380
tatTTTGTATG tattgcaaaa	atagataata	atttatataa	cagttttct	gtttatagat	1440
tggttcaaga ttgtttgg	ttattgttcc	tgtaaagaaa	acaataataa	aaagcttacc	1500
tacataaaaaa aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1526

<210> 34
<211> 1737
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1674)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1731)
<223> n equals a,t,g, or c

<400> 34					
gtcgacccac gcgtccgccc	acgcgtccgc	ccacgcgtcc	ggtttataaa	cagaagttt	60
aacttgcgt	cttaagcttc	cgtttataaa	cagaagttt	aaatttatagg	120
cattcagctc tgtaactca	ctcatcttt	tgttttttca	cactttgtca	agatttctt	180
acatattcat caatgtctga	agaagttact	tatgcagatc	ttcaatttcca	gaactccagt	240
gagatggaaa aaatcccaga	aattggcaaa	tttggggaaa	aagcacctcc	agctccctt	300
catgtatggc gtccagcgc	cttgcattttc	actttttctgt	gccttctgtt	gctcatttgg	360
ttggggagtct tggcaagcat	gtttcatgt	acttttgaaga	tagaaatgaa	aaaaatgaac	420
aaactacaaa acatcgtga	agagctccag	agaaatattt	ctctacaact	gatgagtaac	480
atgaatatct ccaacaagat	caggaaccc	tccaccacac	tgcaaacaat	agccacccaa	540
ttatgtcg	agttatata	caaagaacaa	gagcacaat	gtaagcttgc	600
tggatttggc ataaggacag	ctgttattt	ctaaatgtat	atgtccaaac	atggcaggag	660
agtaaaatgg cctgtgtgc	tcagaatgcc	agccgttgc	agataaaacaa	aaaaatgca	720
ttggaaatca taaaatccca	gagtagatca	tatgactatt	ggctgggatt	atctcctgaa	780
gaagatccca ctgcgtt	gagagtggat	aatataatca	actccctgc	ctgggttata	840
agaaaacgcac ctgactt	aaaatgtat	tgtggatata	taaaatagact	atatgttcaa	900
tattatctc gcactt	aaaaaaatgt	atatgtgaga	agatggccaa	tccagtgca	960
cttgcgttcta catatttt	ggaggcatga	ggcatcaatc	aaatacatttgc	aggagtgt	1020
kggggtgggg gttctaggct	atagttat	ttaaaatattt	tctgggttgc	atttagtt	1080
gtttgtctga agacctggg	ttttatcatg	cagatggaaac	atccaggtag	caagcttca	1140
agagaataga ctgtgaatgt	taatgccaga	gaggatataat	gaagcatgtc	ccacccccc	1200
ctttccatca tggcgtgaac	cctggaggaa	gaggatgtcc	attcagatag	tgtggggggc	1260
cttcgaattt tcattttcat	ttacgttctt	cccccttctgg	ccagatgg	ccagaggcaa	1320

cataaaaaac	cagcaaattt	taatttgtc	ccacagcggt	gctagggtgg	catggctccc	1380
catctcggtt	ccatcctata	cttccatggg	actccctatg	gctgaaggcc	ttatgagtca	1440
aaggacttat	agccaattga	tttgtctagg	ccaggttaaa	atggatatgg	acatgcatt	1500
attacctttt	aaaattatttta	ttttaagtaa	aagccaaataa	acaaaaacgaa	aaaggcaagt	1560
tacggagactg	actttttttt	aacttctgt	tggttgagcta	ctgtaaacctt	ggctttttgtt	1620
aaagacatac	agcaatttagc	tatgcaaaaca	taagcatttt	tctgtaaaaaa	aatntataga	1680
tagatatctt	tatctccat	aactcataac	ttggggatgtat	tataccccgg	ngggcttt	1737

<210> 35
<211> 2242
<212> DNA
<213> *Homo sapiens*

<400>	35					
tcgaccaccacg	cgtccgggct	gccatggcg	cgccggggccg	gctcccgagc	tccctggggccc	60
tcttctcgcc	gtccctcgca	gggcttgcac	tactgggagt	cgccgcggtc	ccagcgcggg	120
cgctgcacaa	cgtcacggcc	gagcttttg	gggcccgggc	ctggggcacc	cttgcggcgtt	180
tcggggacct	caactccgac	aagcagacgg	atctcttcgt	gctgcgggaa	agaaaatgact	240
taatcgtctt	tttggcagac	cagaatgcac	cctattttaa	acccaaatga	aaggatcttt	300
tcaagaatca	cagtgcattt	ataacaagt	tagtccctgg	ggatttatgt	ggagattctc	360
aattgtatgt	ccttgcata	tatcttccca	aaaattatgc	caagagtga	ttaggagctg	420
ttatcttcgtt	gggacaaaat	caaacattag	atcctaaca	tatgaccata	ctcaatagga	480
ctttcaaga	tgagccacta	attatggatt	tcataatgt	tctaattctt	gatatttttg	540
gtatcacaaa	tgaatccaac	cagccacaga	tactattagg	agggaaattt	tcatggcattc	600
cagcattgac	cactacaagt	aaaatgcgaa	ttccacatc	tcatgcattt	attgatctga	660
ctgaagattt	tacaggagat	ttattctgt	cgacattgaa	tgccaccact	agtacattcc	720
agtttgaat	atggggaaaat	ttggatggaa	acttytstgw	magtacymta	ttggaaaaaac	780
ctcaaaaat	gatgggtggtt	ggacagtcag	catttgcaga	ctttgatgg	gatggacaca	840
tggatctt	actgcgcaggc	tgtgaagata	aaaattgc	aaagagtacc	atctacttag	900
tgagatctgg	gatgaagcag	tgggttccag	tcctaaaga	tttcagcaat	aagggcacac	960
tctgggggtt	tgtgcattt	gttggatgaa	agcaaccaac	tggaaatacc	atccaaat	1020
cccttcata	tggagactac	aatatggat	gctatccaga	cgctctggc	atactaaaga	1080
acacatctgg	aagcaaccag	caggccttt	tactggagaa	cgtccccgt	aataatgc	1140
gctgtgaaga	ggcgcgctga	atgtttaaag	tctactgg	gctgacagac	ctaaatcaaa	1200
ttaaggatgc	catgggttgc	acccctttt	acatttacga	agatggaaatc	ttggacattt	1260
tagtgcata	taaaggatata	acaaaagaat	attttgcatt	tcatacacta	aaaaataact	1320
ttgaagcaga	tgcttatttt	gtttaaagtt	ttgttcttag	tggctctgt	tctaatgact	1380
gtcctcgtaa	gataacaccc	tttggagtga	atcaaccctt	accttata	atgtatcacaa	1440
ctgttagatgc	aaatgggt	ctgaaaat	gatcagctgg	ccaactcagc	caatccgcac	1500
attagtctt	ccaaacttca	tacaacgtgc	tgggtttag	tggagcgca	attttcttgc	1560
accatctcta	cgttgtatt	ccccgtccat	ctggagaaaa	atctatacga	aaacaagagt	1620
ggactgcaat	cattccaaat	tcccagctaa	ttgtcatttcc	ataccctcac	aatgtccctc	1680
gaagttggag	tgccaaactg	tatcttacac	caagtaat	tgttctgtt	actgctatag	1740
ctctcatcg	tgtctgtt	ttcatcttgg	caataatttgg	cattttacat	tggcaggaaa	1800
agaaaagcaga	tgatagagaa	aaacgacaag	aagcccaccc	gtttcattt	gatgctatgt	1860
gacttgcctt	taatattaca	taatggaaat	gctgttca	tgattatgt	aaacacaaaat	1920
tctggcttga	aaaaataggg	gagattaaat	attattttata	aatgtatgt	cccatggtaa	1980
ttatggaaa	gtattcaat	aaatatggtt	tgaatatgtc	acaaggatctt	ttttttttaaa	2040
gcacttggta	taaaaaatt	tgggttctt	attctgtatgt	gctgtacatt	tttggccctt	2100
tgtggaaatgt	gttgcattgt	ctccagttt	tgtgtattt	taatcttatt	tgcatcatga	2160
tgtatggaaa	agttgtgtaa	ataaaaataa	ttaaatgagc	aggaaaaaaaa	aaaaaaaaaa	2220
aaaaaaaaaa	aaggccggcc	gc				2242

<210> 36
<211> 2235
<212> DNA

<213> Homo sapiens

<400> 36

gtaattcgc acgagggttc	caccaacatg gagctctcg	agatgtcgsa gctcatgggg	60
ctgtcggtgt tgcttggct	gctggccctg atggcgcacgg	cggcggtasc gcgggggtgg	120
ctgcgcgcgg gggaggagag	gagcggccgg cccgcctgcc	aaaaagcaaa tggatttcca	180
cctgacaaat ctcgggatc	caagaagcag aaacaatatac	agcggattcg gaaggagaag	240
cctcaacaac acaacttcac	ccaccgcctc ctggctgcag	ctctgaagag ccacagcggg	300
aacatatctt gcatggactt	tagcagcaat ggcaaatacc	tggctacctg tgcatgat	360
cgcacccatcc gcatctggag	caccaaggac ttccctgc	gagagcaccg cagcatgaga	420
gccaacgtgg agctggacca	cgacacccctg gtgcgttca	gcccgtactg cagagccttc	480
atcgcttgcg tggccaacgg	ggacacccctc cgtgtttca	agatgaccaa gcgggaggat	540
gggggttaca ctttcacagc	caccccgagag gactttctta	aaaagcacaa ggcgcctgtc	600
atcgacattt gcatgtctaa	cacagggaag tttatcatga	ctgcctccag tgacaccact	660
gtcctcatct ggagcctgaa	gggtcaagtg ctgtctacca	tcaacaccaa ccagatgaac	720
aacacacacg ctgctgtatc	tccctgtggc agattttag	cctcgtgtgg cttcaccccc	780
gatgtgaagg ttggaaagt	ctgctttgga aagaaggggg	agttccagga ggtgtgcga	840
gccttcgaac taaagggcca	ctccgcggt gtgcactcgt	ttgtttctc caacgactca	900
cgaggatgg ctctgtctc	caaggatggt acatggaaac	tgtggacac aratgtggaa	960
tacaagaaga agcaggaccc	ctacttgc	aagacaggcc gcttgaaga ggcggcggt	1020
gcccggccgt gccgcctggc	cctctccccc aacgcggcagg	tctggccctt ggccagtggc	1080
agttagtattc atctctacaa	tacccggcgg ggcgagaagg	aggagtctt tgagcgggtc	1140
catggcaggt gtagcgc	tttgccctt gacatctg	gcccgtttctt ggcctctgt	1200
ggggaccggg cggtcggt	tttgcacaac actccctggc	accgagccat ggtggaggag	1260
atgcaggggcc acctgaagcg	ggcctccaaac gagacaccc	ccagagggct gcagcagcag	1320
ctgaccagg cccaagagac	cctgaagagc ctgggtgccc	tgaagaagtg actctggag	1380
ggcccccgcg agaggattga	ggaggagggg tctggctcc	tcatggcact gtcgcacat	1440
ttccctccag gtggaagcct	ttcagaaggaa gtcctctgg	tttyttactg gtggccctgc	1500
ttcttccat taaaactact	tttgtctact taggtcttcc	tcttcttgc ggctgtgact	1560
cctccctgac tagtggcaa	ggtgcttttcc ttccctccag	gcccagtggg tggaatctgt	1620
ccccacctgg cactgaggag	aatggtagag aggagaggag	agagagagag aatgtgattt	1680
ttggccctgt ggagcacaat	cctcacaacaa aaaaagattt	gtaaatgttc cagaacaacc	1740
tagagaacac ctgagacta	agcagcaggat ttgcaaggat	gggagactgg gatagtttcc	1800
catcacagaa ctgtgttcca	tcaaaaagac actaagggtt	ttccctctgg gcctcagttc	1860
tatggagaata atccctctcg	tgaactctt gcaaaagatga	tatgggcta tatggctcaa	1920
agagaatatac aagtccccag	gtctggaa aaaaatggaaa	agagtagtac tattgtccaa	1980
tgtcatgaaa gtggtaaaag	tgggaaccag tggctttga	aaccaaatta gaaacacatt	2040
ccttggaaag gcaaaagttt	ctggacttg atcatacatt	ttatatggtt gggacttctc	2100
tcttcgggg atgatatctt	gtttaaggag acctctttc	agttcatcaa gttcatcaga	2160
tatggatgt cccactctgt	gccccaaataa atatgagctg	gggattaaaa aaaaaaaaaa	2220
aaaaaaaaaa ctgc			2235

<210> 37

<211> 2971

<212> DNA

<213> Homo sapiens

<400> 37

gacgtgagga gcgttccatt	tggccagttt tggccgggtt	ccacagctgg tttagggccc	60
cgaccactgg gcccccttgt	caggaggaga cagectcccg	gcccggggag gacaagtgc	120
tgccacctt ggctgcgcac	gtgattccctt gggacggctt	gttcctgc gtcagctgc	180
ggccgagttt ggtctccgt	gttcaggccg gtcctccctt	ccttgcgttcc ttctcccg	240
tggccgggtt tateggagg	agattgtctt ccaggctag	caattggact tttatgtat	300
tttgaccagg cggcaggaat	agcaggcaac gtgatttca	agctggctc agcctctgtt	360
tcttcctctcg tgaatcgca	aaaccattt tggagcagga	attccaatca tgcgtgtat	420
ggtggtgaga aagaagggtg	cacggaaatg ggagaaactc	ccaggcagga acacccctt	480
ctgtgatggc cgctcatga	tggccggca aaaggattt	ttctacactga cccttttct	540

catcctgggg acatgtacac tcttcttcgc ctttgagtgc cgctacctgg ctgttcagct	600
gtctccgtcc atccctgtat ttgctgccc gctcttcctt ttctccatgg ctacactgtt	660
gaggaccgc ttcaatgtacc ctggagtgtat ttctccggcg ctaccatgt aagcagctt	720
catagaaatg gagatagaag ctacaatgg tgccgtgccc cagggccagc gaccaccgccc	780
tcgttatcaag aatttccaga taaaacaacca gattgtaaaa ctgaaaatact gttacacatg	840
caagatcttc cggcctcccc gggcctccca ttgcagcata ttttgacact gtgtggagcg	900
cttcgaccat cactgcctt ggggtggggaa ttgtgttggaa aagaggaact accgctactt	960
ctaccccttc atcccttctc tctccctctt cacaatctat gtcttcgcct tcaacatctgt	1020
ctatgtggcc ctcaaatactt tgaaaattgg cttcttggag acattgaaaag aaactccctgg	1080
aactgttcta gaagtctca ttgttttt tacactctgg tccgtcggtt gactgactgg	1140
atttcataact ttcttcgtgg ctctcaacca gacaaccaat gaagacatca aaggatcatg	1200
gacagggaaag aatcgcggtcc agaatccca cagccatggc aatattgtga agaactgctg	1260
tgaagtgtt tttttttttt tggcccccag tttttttttt tttttttttt tttttttttt	1320
ggaggaaaatg ggaaggatcgac ctcccgatgt tttttttttt tttttttttt tttttttttt	1380
gagcccgccccc cccacagaaac acctgaactc aaatgagatg cccggaggaca gcagcactcc	1440
cgaagagatg ccacccctccag agccccccaga gccaccacag gaggcagctg aagctgagaa	1500
gtacgcctatc tatggaaagag actttttttt tttttttttt tttttttttt tttttttttt	1560
gtgagaagtt aaacctgaga cagagagaaaa gtaagctgtc cttttttttt tttttttttt	1620
ggtttttagt caccctgggtt cacactggca tttttttttt tttttttttt tttttttttt	1680
gaactcaagg cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1740
tggcccttgg cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1800
tccctccaga tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1860
ttttagactg gttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1920
gagaccccttgc gttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1980
gagactgggg tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2040
taggaatgggg acaggtaacct tttttttttt tttttttttt tttttttttt tttttttttt	2100
cccatccact ctgacacacta agccccactc tttttttttt tttttttttt tttttttttt	2160
agtagagata ataatttgaca tttttttttt tttttttttt tttttttttt tttttttttt	2220
cattttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2280
ctctttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2340
taggacggg aatggcgatg tttttttttt tttttttttt tttttttttt tttttttttt	2400
agcctggatg tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2460
cctgaatgca caaatggaa accaaggcac agagaggctc tttttttttt tttttttttt	2520
cgatgtaccc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2580
tgagggaaag cccagcactg tttttttttt tttttttttt tttttttttt tttttttttt	2640
tctggctatg gtaaccacac tttttttttt tttttttttt tttttttttt tttttttttt	2700
cgccagagtc ccagagccac tttttttttt tttttttttt tttttttttt tttttttttt	2760
tcaggacactg tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2820
gagccactgt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2880
ggagtgtaaa ccggaaactt tttttttttt tttttttttt tttttttttt tttttttttt	2940
aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa a	2971

<210> 38
<211> 1163
<212> DNA
<213> Homo sapiens

<400> 38	
ccacgcgtcc gccaagggtt ctgatttaggg aatggttatg gacttagggat atcagtaaca	60
atggtttagaa atggctaac atttggatgg cacctgtgt gtgcctggcc cccgctggga	120
gccttcgtgc ccagagtgtac cccgtctgaa tttttttttt tttttttttt tttttttttt	180
gtggggggca gagctgcaca actcacatgt gccgagctca agactcactc ctgggtctgc	240
ctgggtctgg ctgtgttgc tttttttttt tttttttttt tttttttttt tttttttttt	300
ccaggatccg cagaacgttc cccggaggagg tttttttttt tttttttttt tttttttttt	360
ctaaaaaggt gatagagtta cactggagag agcagcatcc aggtgcagca gggatggggcc	420
tggggctcac gggcaggggc tttttttttt tttttttttt tttttttttt tttttttttt	480
tgtgtcagaa gcactcgatg aatctttttt tttttttttt tttttttttt tttttttttt	540

gcttgctgtc	gcattgcctg	caatccctgga	tgaatgccca	ggttggcttt	gtcccccgtc	600
gggtggatgt	gacgttagct	gtgatgttag	gtccctggct	ttaaaatacg	acggaaactgg	660
gaattggagg	agcagtgggg	gcagaaaagg	cagccccgca	gaggcctgga	gctgagcagt	720
gcgggcgacc	caggagcgt	gagtgcctcc	gtcacagcct	tcatcgacc	ctgtggctct	780
cataaagggg	atggaatcta	cgaatttagt	tttcccagcc	tccttaaaaa	ctattcatg	840
ccaggggcag	tggctcacac	ctgaaatccc	accacttgg	gaggctgagg	caggctgatt	900
acttgaggtc	aggagtttg	gaccagccta	gccaacatgg	tgaaaccccg	tgtctactca	960
aagtacaaaa	aaaaaaaatta	gtcagacgtg	gtgtcacgca	cctgtaatcc	cagcttttg	1020
ggaggctgag	gcaggagaat	cacttgaacc	caggaggcag	aggttatagt	gagccagtat	1080
tgcgccactg	acctccatct	gggcaataga	gtgagaccct	gtctcaaaaa	aaaaaaaaaa	1140
aaaaaaaaaa	aaaaaaaaaa	aaa				1163

<210> 39
<211> 1932
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1624)
<223> n equals a,t,g, or c

<400> 39						
ggcacgagcc	aggccctgg	gccgggcgt	gaggcgggca	cctctggca	gggccccggc	60
ggggctgggt	ggccggcccc	tgctgctgcc	gtccatgctg	atgtttgcgg	tgatcggtggc	120
ctccagcggg	ctgctgctca	tgatcgacg	gggcattctg	gccgaratga	agccccgtcc	180
cctgcacccg	cccgcccg	arggcacagc	ctggcgggg	aaagccccc	agcctgggg	240
cctgtccctc	agggtgggg	acgcggactt	gcaagtgcgg	caggacgtcc	gaaacaggac	300
cctgcggg	gtgtgcggac	agccaggat	gccccggac	ccctggact	tgcgggtggg	360
gcagccgcgc	accctgctc	gcccattct	cgtaagtgc	cgttaccgc	tctctactcg	420
ctacgtcccc	aagggtggct	gtcttaact	gaaaggggt	atgaagggtc	tggcaggcgt	480
cctggacagc	gtggacgtcc	gcctcaagat	ggaccaccgc	atgtacctgg	tgttccctggc	540
cgacctgccc	cctgaggaga	ttcgctacc	cctgcagcac	tacttaagt	tctgtttgt	600
gcgggagccc	tttggAACGCC	tcctctctgc	ctaccgcaac	aagttggcg	agatccgaga	660
gtaccagcaa	cgctatgggg	ctgagatagt	gaggcgtac	agggctggag	cggggcccag	720
ccctgcaggc	gacgtgtca	cattcccca	gttccctgaga	tacctgggt	atgaggacc	780
tgagcgcatg	aatgagcatt	ggatggccgt	gtaccacct	tgcagccct	gtgcgggtca	840
ctatgactt	gtgggctct	atgagaggct	ggaggctgtat	gcaaatcagg	tgctggagtg	900
ggtacgggc	ccacccatc	tccgattttc	agctcggcc	gcctggtacc	ggccagccag	960
ccccggaaagc	ctgcattacc	acttgtgc	ggggccccc	gcctgcgtct	aggatgtgt	1020
gcctaagtat	atccctggayt	tytccctt	tgccatccca	ctgccta	tccacaagg	1080
ggcggtcgt	cagtgaccat	gggtgtgggg	ccagcagct	gtggggact	gtttcaacgc	1140
cagcttctg	tgcttctgcc	tgtcattcgg	agaaactct	gctctgggg	ttggggcttc	1200
tcaggatct	ggatggcaga	gactgcctc	agaarttct	tgtccagggt	gggcacccac	1260
agtgactc	aggacagggc	taggcaggag	acctgtct	cctcatttgg	gggatctct	1320
ggggggcaga	caccatgtt	ccaatgaagc	aacacatct	atctaaagac	tggctccaga	1380
ccccggctg	ccaggattat	gcagtcac	tggtctac	taatttacc	tgtggccaa	1440
ctcagagat	gtaccagcc	ggggcaagca	tgaccagac	cagggacc	gtggctct	1500
tccccat	atccacccca	tgtgcctc	gactagatg	agaatcata	ccttataat	1560
gactttgt	ccttcgtct	ccagtct	aatttcctac	acctgccc	tttttacatt	1620
tttnccaagg	aaaggaaaac	ggaagcagg	ttcttgcct	gtagctcc	gaccctc	1680
tgcaggcacc	caaaagacc	ctgtgcctc	gaggctct	aaacccct	c	1740
ctaattctcc	cttccttccc	cacaaggcm	ttgagg	ttgtgtgg	gttatatct	1800
gctgcattt	ttctgtat	tttattt	atttgactt	tttgatagaa	cccttgtaa	1860
ggcttgttt	tcctaatagc	tgactttt	ataaaagc	gttataaaa	aaaaaaaaaa	1920
aaaaaaaaaa	aa					1932

<210> 40
<211> 881
<212> DNA
<213> Homo sapiens

<400> 40

gaattcggca	cgagggaaacc	cagaagatgc	tgcctctcct	gatcatctgt	ctcctgcctg	60
ccatttgaagg	gaagaactgc	ctccgctgct	ggccagaact	gtctgccttg	atagactatg	120
acctgcagat	cctctgggtg	accccaggc	cacccacaga	actttctcaa	agtattcact	180
ccttggcct	agaggataat	aattttctca	aaccctggta	ccttgatcgt	gaccatTTgg	240
aagaagaaa	acccaaattc	ttcactcaag	tacaccaagc	cattaaaacg	ttacgagatg	300
ataaaaaacat	acttctggaa	gagatctaca	cgcacaagaa	tctcttact	gagaggctga	360
ataagatata	tgatgggctg	aaggagaagg	gagccccacc	cytctccatg	aatgccttcc	420
cggtccatc	tcttacttgc	accccagaac	cccttggctc	tgtctgcctc	cccagcacct	480
cagtttctct	accttctcac	ctccctggca	gcctgcaatg	agtctctgtgc	caggaacccg	540
cggaaccccc	tgtgggctgt	gagtctcagc	agtgtctac	tcctggccat	agctggagat	600
gttttttta	ctggcaaagg	aagaaggagg	cagtaaagga	acagggcage	ccgcatgtct	660
tccagaagtg	aacagaggcc	gcagctacca	ccgtcacaaa	gttcaactcat	ctctgggtcc	720
cggtgacccc	atccccccat	accctccatc	ctgggtccctg	ggggcccaaa	gctctgaggg	780
ctaggagact	gcgtgtctc	gtgggttggc	tactctaca	cctttgtaaa	gagtctcttc	840
attaaaaacc	ctcttcataa	aaaaaaaaaa	aaaaaaaaactcg	a		881

<210> 41
<211> 1932
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (2)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1022)
<223> n equals a,t,g, or c

<400> 41

cncggcgcgg	ctcggctcat	gccccccggc	gcggggcaca	caggccggcc	ggcagccgct	60
ggggaaatagg	ccccccgggg	cggtgtccgc	ggcggggcca	tggcggag	accccccggc	120
ccggccgcct	ccggggagga	gttctccctc	gtcagccgc	tggtaaaata	cctgtctttc	180
ttcttcaaca	tgctttctg	ggtgattttc	atggtgatgg	tggtgtgtgg	tgtctacgct	240
cggctaatga	agcatgcaga	agcagcccta	gcctgcctgg	cagtggaccc	tgccatcctg	300
ctgatctgtgg	tggtgtctt	catgttcctg	ctcacccct	gtggctgcat	tggtcccttc	360
cgcgagaaca	tctgcctct	gcagacgttc	tccctctgcc	tcacccgtgt	gttctctgt	420
cagctggccg	ctgggatct	gggcttcgtc	ttctcagaca	aggctcgagg	gaaagtgagt	480
gagatcatca	acaatgccat	tgtgcactac	cgagatgact	tggatctgca	gaacctcatt	540
gattttggcc	agaaaaaagt	tagctgtgt	ggagggattt	cctacaagga	ctggctctcag	600
aacatgtatt	tcaactgctc	agaagacaac	cccagtcag	agcgtctc	tgtgccttac	660
tcctgttgct	tgctacttcc	tgaccaggca	gtgatcaaca	statgtgtgg	caaaggatgt	720
caggccttgc	actacttgg	agcttagaaa	gtcatctaca	ccaatggctg	tattgacaag	780
ttggtcaact	ggatacacag	caacctat	ttacttggtg	gtgtggctct	aggccctggcc	840
atccccccagc	tggtggaaat	tctgtgttcc	cagatcttag	tgaatcagat	caaagatcag	900
atcaagctac	agctctacaa	ccagcagcac	cgggctgacc	catggtactg	agaatccatc	960
ctgcacccctcc	tcaccatgga	aactggcaag	cctcataaac	gaacagcagt	gggtgctgaa	1020
ancagcacca	aatggagatt	tggattccag	ccccccactg	acagcccagt	ggaaagaagc	1080

aaactccaga tgggcagaag gcagggtgca caggtggctc cagtctcagg aggatgcgcc	1140
tcctctcccc catcccagcc ctcagcattg tgccagagt atacccttaa gtgtttgggt	1200
ttatttttc agttttgtt gggaaacagc agttcacag agatgtgggg gtactgctgc	1260
tgcctttca ccgaggact gccaccacca gctctascag ggatgctcct gagcttggcg	1320
gacatactta gatcctaacg tgccagttag acctggctgt ggagagtagc actggcagcc	1380
ctgcctggac tccacttggc atgataccag ctccagaagg gaaggagtg gaggcaggcag	1440
tgaggagaga gcctgggggt cggctggga cagccgtatg tgcttaggtag gagtggaggg	1500
agatatgtt accaaatgcc tgcctgca tcctccagg tagtcagagt gagctacatc	1560
ctgccccgcc ttcatttcca tggaaacatg gcagcttagga cacgggttac aacagcagcc	1620
aaatcttcc ccaacctccct tacttcgaaa aaaagttgg aaccctggc cctatactct	1680
gcagtcagaa gtgggactgt gccatacatg cccttgaatt cctccctgtc tggccctccc	1740
tctccagcaa gcagggtttt cttaacttg gcagttgtca gaggagaagt ggtAACACCC	1800
ccaccccccatt cccctgcattt ggagctcagt attccatag ggttaagaggt aggaatcttg	1860
ctggacgag gggagccaga agtggcaata aaagctgtt gacctggaaa aaaaaaaaaaa	1920
aaggcggcc gc	1932

<210> 42
<211> 1164
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (582)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (592)
<223> n equals a,t,g, or c

<400> 42	
ggcacgagct tgggtgtcac cagccctctg atctgccagg gtctgctctg ggttggact	60
gaccagggtg tcatcgctt gctggccctg cctcggtgg aaggcatccc caagatcaca	120
gggaaaggca tggctctact caatggcac tggccctctg tggccctctt ggctgtggct	180
accagcatcc tggccctctg catcctgcgg agtgaccagg aggaggctga gggggcccccgg	240
gctgaggagg caaaggccaga cgggcaggca cacgagccca tgccctggcag ccacgtggc	300
cgagagctga cccgcagaa gggcatctc ttgcagttacc gcctgcgtc caccgcacac	360
ctccggggcc cgctgtctc catcgggag cccggccctg ctgtggcgc agctttggag	420
cacagcgagg aggacggctc catttacgatggccagc accccgacgt ctgggtgcgc	480
agccggccct gcccggcga cggccaccgc aaggagattt gctctgtggc catcatctcg	540
gcggggcaggg ctacccgcaac ttggcagcg ctctggccac antgggaagc angccccgtg	600
tggggagacg gacagcaccc tcctcatctg gcagtgccct tgatgtata ggcctccccc	660
tctccctca gagggcacag ctgcaggctt gaccaaggcc acggccggct ctcgtgtct	720
aggacctgca cgggacttgtt ggatggccct ggactctcca gaaactactt gggccagagc	780
aaagaaaaac ctcttggttt aaaaaaaaaaattt ttccagagt gttttggggg ggagtttttag	840
ggcttggggg gaggggaggac acatctggag gaaatggccct tctttttaaa agcaaaaaaaac	900
acaaaaacctc acaactgcct ggcaaggccca gtatcttgc ttggggccct agcgggactc	960
caaggcagcc acacggccctt cctggaaaggg tgggtgtcggt tgatgtgtg cgagtgtgt	1020
ggcttgggtg tgaatatcta taaaataatgat tatatgtgtt atatatatgt tggataaaaata	1080
aagtctgtac atattggagc tctggagat gctggaaataa aagacaagag ttacatctgg	1140
acttggaaaa aaaaaaaaaaa aaaa	1164

<210> 43
<211> 1105
<212> DNA

<213> Homo sapiens

<400> 43

gaattccggca	cgagaacaaa	ttgaaaccat	ctggcatga	acttttattt	gttaagagg	60
tttctaatat	tgattcaatc	tctttgctta	ttctaaatgt	gttcagat	tccacttctt	120
gagtcaattt	ggtat	ttgttctag	gaatttgtcc	atttcata	gtttaccaa	180
tttttgcacat	ataaaattat	atatggaaat	ttctaaaata	tttaaaaatt	tctgtatgt	240
caatagtaat	gtcccctctt	ttgttaccaa	ttgttattt	gaatcttc	ctttttttg	300
tcaatctgc	taaaaatttg	tcaatttgt	tcgtctctc	aaaaaaat	acgtttgct	360
tcatgattt	tctawtgtt	ttccatcyat	atttcattt	aatacattt	taaacytay	420
cttttattt	tcatcccttc	ttggagactt	gggtcttatt	ttttttcc	gataatctag	480
ttgttattt	tataagatta	agtatttt	tgaaatctgt	atgttcttta	atgtaggcat	540
tcaactactat	aaatattactt	ctcaggagca	tctctccgc	attccatgtt	tttagtatgtt	600
gtgtttaat	ttgttattcat	aactagaggg	aaacagagg	gacggagaaa	aagacgtaca	660
aatatcatcc	acttgcaag	tatagattt	tttgatgtk	rataatgatr	aaaatattac	720
gagacagata	agaaaatttg	aacactgacc	attgatgcag	ttacagttaa	ttttaaaatc	780
aaggtaata	acattttagt	tatTTaaag	aatgatagta	atttagagat	gtattctgaa	840
tgttttaaa	tgaaaagata	tgccctggat	ttcttccaa	atgaatctt	taggttggg	900
agaaaatgag	aacatagtgg	aaacaagact	gacaatgagt	tgttgagg	gggcaatgcg	960
tacactaaag	cttattttat	cttattttac	tgtatatact	gttaaagctt	gcattatttt	1020
cataaatgca	tttgctaa	gcaactgtt	tcaaataaag	tggattggc	tctaaaaaaaaa	1080
aaaaaaaaaa	aaactcgagg	ggggg				1105

<210> 44

<211> 1262

<212> DNA

<213> Homo sapiens

<400> 44

cagcatgtac	ccagttgttc	tttctcctga	gaaagaaaa	tgctctgat	ttcttataat	60
ccaggctgcc	acgtttacct	tgtaaaatca	atacttaatt	tttagatttt	tatattatct	120
tttctcgtga	agcaagactt	ctaaattatg	gctataat	ctttgaatt	tttgttctt	180
atgaatcttc	caactgtaaa	ctcatcta	ttcaaaactt	tcatacctg	ggatgtaa	240
ttgtccttgc	tttctcatct	tgatattacc	gtcaatcatt	ttgttattct	gagtagattt	300
gaacttgctg	gagtaataga	ggggaaaccc	ctgectgtt	ctaaatcaga	tctttgtcct	360
atactcggac	aattatggtt	tcatatttta	ttatTTtttta	tttctgggt	ttaacaaatg	420
agataacat	tttagacataa	tatTTgtaaa	catcttgact	tatttcagca	ttttcccttt	480
ttgtgtatct	tcagagat	ttgttggaa	agcaatttcc	aaatTTttt	aaattattga	540
agtctactag	cacggaaaggt	caaattctt	ggatattaa	aaaatgtt	ttaataatca	600
aactcatctt	aaaaaaatgtt	catcagactt	tgttggat	gcacatttt	ccaaaagaga	660
gccttatttc	tgtggaaagaa	atacagtat	tactttgg	tttactaa	aaaaactgtt	720
actttaaggc	acagagcaga	tatagaatcc	cccttcttcc	ccactcctag	tgactggat	780
tctacattaa	tatTTatctt	ccatgcata	tgtacttgag	ggaaaaaaac	aataactctt	840
aattgtttaa	tatcaaacaa	taaaatctg	tgtatctgt	actgtcaata	gatggctt	900
tgtttaaaaa	ctgaagctac	tccagaagta	ggaattaa	tatTTgtt	acaaagtca	960
tcaaaccaga	gccatgtcct	ggggaaactgt	caaaagaat	gttccatagg	gccagaggcc	1020
acatccactg	ttagatgaca	gaacaaccat	acttcagat	gcaaaaccgg	tcagtttgg	1080
ttgcgttgc	tgccatctt	ctttctgtt	gtctcagct	aattaatgtc	tggagagct	1140
caaataatgtt	aaatagatcca	agatgacca	ttctgcagg	tggcaaggct	gatcttgc	1200
ttttgattaa	aataaagaac	attccccaa	aacagttt	tgcaaaaaaaaa	aaaaaaaaaa	1260
aa						1262

<210> 45

<211> 517

<212> DNA

<213> Homo sapiens

<400> 45	
gaattcggca cgagtgcact tccaccagct atgtatgaga cttcccattt gtcacatct	60
ccagttttt atgtggtcag tccttttgc ttgggtcatt ttgggtgata tgaaatggca	120
tctcatgttg gcttttcat atatccctt gatgactaat ggtattttt cacccttca	180
gtgcttattt gccatttcatg tatctttttt ttgtgttag cacttcagg ttttgc	240
tagatttatg gggttgattt ctctttatata atgatggta gggatgttat atatattctg	300
gacacaagat tattttaga gatacgtact tcagatattt tctcccgatc tgtagctgc	360
ctaatttattt ttatttattt tattttagat gaagtctcac tctgtcgccc aggcagaggt	420
tgcagtggc cgagatagca ccactacact caagcctggc tgacagagtg agactctgtc	480
tcaaaaaaaaaaaaaaaa aaaaaaaaaaaa aactcga	517
<210> 46	
<211> 858	
<212> DNA	
<213> Homo sapiens	
<400> 46	
aaaaaaaatc ctacatggat attggtagga aagagagaaa ggaagtggcc agtgtcccg	60
ggccttcc accttcttgc ttgttgaagc tggggcctgg agggatgg cctggccactc	120
agcagggggc actaatggga ccaagctaac ctgttcagtg agaatcctgc agggagacct	180
gagggttacca gggaaagtgc ggggaaaggcc cgggaaatgg agagagctgg tctgggggg	240
aggagcaagc cgcgtggggc agggccatgtg cttttgcctt gggggatgt tactcattt	300
gagctgtgcg tcttggAACgc ctgcctcaca cacaaggac tggggcagat gtaagttctc	360
tgcagcaacg aagcgcacag ctgagagtaa cttagaaagc acccagctaa tgctggcatc	420
ccagatcgac cccctctcg ctgaatgtt gcattctctgt gcctcagttt cctcatctgt	480
aaatgggggt gataagaaat gtgtacacac ctccgggca gtggggagga taaaactgtg	540
ctctgacacg atccgggcat gttcagggtg gtatctgcag taaaccgcgc tcggaaaatg	600
gcggcgcatac agggccagcg gtgggagctc tccgtgctt gcttgacgcc attgtggagg	660
tggaggaggg gctgcaagac tctgagcagg aagccccgc aaagcaggaa agcagagcca	720
gatgtggggg ccagccgcag aaacgagagc cccctgtact ttgaggcacc ctggagag	780
ggcagaaagc aggaagggtt aattttctcc aaaacccaag aggagatgt accccacatg	840
ataactgagt ttctcgag	858
<210> 47	
<211> 6107	
<212> DNA	
<213> Homo sapiens	
<220>	
<221> SITE	
<222> (5749)	
<223> n equals a,t,g, or c	
<220>	
<221> SITE	
<222> (5892)	
<223> n equals a,t,g, or c	
<220>	
<221> SITE	
<222> (5896)	
<223> n equals a,t,g, or c	
<220>	
<221> SITE	

<222> (5906)
 <223> n equals a,t,g, or c

 <220>
 <221> SITE
 <222> (5957)
 <223> n equals a,t,g, or c

 <220>
 <221> SITE
 <222> (5966)
 <223> n equals a,t,g, or c

 <400> 47

gcagtttagtt	ccttgatgtc	agtagtgggc	taaaggcagc	ttactgtgtg	tttgctggag	60
ctttcactca	gccaagtgtt	agagtcagga	aaccattga	ggcaatggcg	tcaaatggtg	120
<tccacaaga	atgagccatt	cagtcttgc	tcactatata	ttaaatattt	tattattgtt	180
gttattgtt	ttattaaattt	gcttctgtt	ttctatgcct	tttatttata	aagacactaa	240
aaaaacccat	gtttgttaatt	ttaataaact	tttcccattc	ttgttaatata	cagagctact	300
ttataaattt	tctgaaccaa	aagtattttc	ctcagtgat	ctcttctccc	ccagcccccta	360
ttggggaaaaa	ttacccagta	tagttcagg	tatgaggagg	atcagccaca	caatccagtg	420
cttcagttt	aaaatgtaaa	attctaaacc	taaagtaggg	ttggttgaaa	tttcagacaa	480
agcaaaaccc	gcaggatataa	aaagtagtat	aaatacaaat	ctgtatgtt	tttttgaatt	540
ttctgaacctt	ttttcttaaga	gattacatag	gagactaaag	aaatctatct	gttcaagttc	600
taatttagat	gattgttaat	actgcactgt	ggatgaatgt	gcgactggct	tgtgtgctga	660
cttctgttgtt	tttagcaagag	gtttattgtt	atcaaatgt	aatttggcaat	gccaagtcac	720
tgggaccaat	tttctgtttt	ataatatctt	agttttagaa	agaatataata	cctgaaactt	780
agtgggttga	tcggatggag	acagaaaaacc	cgattttat	tctcataaat	tttgggttta	840
tttataacaag	ggctgtgcta	tgctaccata	ttcttggtca	ataataatag	gtttgttgg	900
ttttttacat	tgtttaatgt	tccttacccc	taaaggctaa	tgttaagtac	aacattctga	960
aaatacaatt	tggctacgaa	gagtattcat	cttctttgaa	gctcagtggt	tgatattttgt	1020
gctaataatg	caatttcctg	attctgttta	caagttatag	ctacatatgg	gagagactca	1080
gtgagccagc	aaaggccat	gaaacaacaa	tttattaaat	gtatttatgg	cagaaggacc	1140
taaataaaact	gtgagccacc	ttttcttctt	tatattgtt	catttaatgt	ttcttgcctt	1200
cagcaactca	cattaatgt	tggagcttat	ctctttctct	ctctctctct	ctctctctct	1260
ctgtgttgtt	gtgtgtatgt	gtgtgtgtt	gtgtgtgtt	ttcccttatty	tcattccatt	1320
atatatccac	accaacatgg	gtgacgataa	ttcaaaatgt	tattttgcct	ctaagcttga	1380
tcatgttacc	ttttaggatta	aaagtatcatg	ttattttagcc	aatgcaaaatc	tgttttaaaa	1440
caaataagtt	aaaaaaaaaa	caagttttta	agggttttat	tatagaagaa	gtattaatga	1500
aggactttcc	ttcctccctc	cctttctcc	cctccctgcc	tcccttcttc	ccttccatct	1560
ccccctccctc	cctgccttct	ttgtttctcc	ttcccttatt	cctccctccc	tccttctcc	1620
cttccctct	ttttttccatt	cattccttct	tgccttttat	tttttttttt	tgtatattca	1680
catgtgttgt	agtttggaaat	tttatttctag	tgcattttct	gctcatcaga	acctcagcta	1740
atctacccat	gaaaaatagt	atcaaaggaa	atggaaaatg	tgtatctgag	tcctccaggaa	1800
actaaagatata	ttctttttgtt	ccatttaaagc	ctttataaat	cggttttgac	cattaaagcc	1860
tttataaaatg	tttggtttttag	gaaagtgtat	cttggatgt	catcaaaaaa	taatgaccag	1920
gacaaaaacga	ttaataaattt	aaagtctcaa	atcaccatgg	ttatatactt	tcaccagaaa	1980
tagtaatctt	acaatttttc	attttctgt	tgaagatttc	tgttccaata	tctgtttctt	2040
aatagatttt	ttaaaatata	tagcttctt	ctgctttatg	accacaggtt	ttatccctaa	2100
ccgagacagc	tgtcttatat	ctgcatgtcct	tagactgtgt	ggagggactc	catgaagaaa	2160
gaccatagtt	tagaaaaataa	actctatgt	tataccctag	taagtgggtt	agttagaatct	2220
cataacatgt	ataaaaaaaga	ggttttcttc	tctgcttgtt	tgtgtcacta	gagcaaaatt	2280
gttagagat	tgctcataat	gcaagataaa	tcagaataat	ctacaatatac	atttgtggat	2340
ggttcccaat	cccagtgtct	tagttactt	actttttttt	ttttttttgtt	gatggagtct	2400
tgcctgtct	ctcagcttag	agcagtgtgc	gatctcaggt	cactgcagcc	tccacccccc	2460
agttcaagc	gattctccctg	cctcagccctc	ccaagtaggg	aggattacag	gcacccctcca	2520
ctaggccccgg	ctaatttttt	ttgttattttt	tttagtagaga	tggggttttt	ccatgttggc	2580
caggctgggtt	tcgaactctt	aacccctcagt	gatccaccc	cctcggcgctc	ccaaagtgtct	2640

aggattacag	gcatgagcca	ccacatccgg	cctaattact	tcttaatcc	ccatttattt	2700
ttatgcatt	ctagcctcat	ttattaataa	aattatgttt	ttacttctc	tttcaggaaa	2760
tttttaat	taatattta	tatctagatc	taatgtatg	gaaaagtgcc	tttttatcat	2820
ttataattc	attttcact	atttccaaa	acacataaaac	aaatagttt	agtaggtccc	2880
agctttaact	tttccattt	aaacottt	tttcattt	cttcccttgc	gtttaagaat	2940
aaaagaaaaa	gtacattgct	agaattgtt	cttgggaga	gggaaaaga	ttacagaatt	3000
agactgttca	gcctttat	aaactaaatt	tgtcttcate	tcaacccagct	aatggtaggt	3060
cttatctgaa	tactcatgag	aattttagca	tctgtgaaac	tccatgcacc	agatgtgtgt	3120
aaatttcagg	aagaaagtgt	tgaagcatt	ttctctgtat	ttaatttagat	gaaataaaat	3180
cactaaaaca	tagtttagt	aaagcctgat	tatgcacatt	ttttttaact	agacaggcga	3240
aagttttta	tgttagtgt	cttcttgct	atcctcagtt	aatttaccta	gacaaaaagt	3300
gtcaaaaggaa	atgagaaaaa	ggttatatct	gactccctcc	agacctaaga	taattcctt	3360
tgatcagata	cagtcatgt	gagtgcctg	gtttttgtt	atttgcctc	tattccagct	3420
ccttaccaca	gccccgtgt	ttaaagaag	gatcatcagc	aacaggctcg	gatagttcta	3480
ccttggat	aggcgtgtt	tccctgtct	agtattctg	tgactgttag	ttggcactgag	3540
gactgaaac	ttttatgca	tattttat	accctattga	tattatgcac	ttaatcatt	3600
ccaaagaagc	caagaatgct	gtatagtat	gattcctcc	taatgaattc	atcttaacta	3660
tttagaatgt	tatgtccctt	ttcttttgg	tagccaaatt	ggtataaaatg	ttatatggat	3720
ttttctaaa	tgactatata	ggacttaaga	cttggaaatg	taatttactt	ataagggaa	3780
ataattatgc	tttagcacat	cattttagaa	acgtcacatt	ttagaaacat	ttagcttgc	3840
aacctacatg	tttggaaatt	cattaaaacc	agttgtctat	atattttgt	ccatgtat	3900
aagaacatca	caatataatct	ttttctacat	atgttagtgc	tgaacccagt	ggttctcaga	3960
gtatgttct	cagccccacca	gctagatcata	gtatcacctg	ggaactagtt	agaaatgtaa	4020
attcttggc	cccatccccag	acatactgag	tcaagaaattc	tggaaatagg	cccccccaat	4080
ctgttttcac	aagccctcca	ggtgatctg	atgcacactt	taaagtttag	gaaccactgg	4140
gctaagactc	tgttgagata	tagatttt	cttccactca	gactgatata	gttatacatt	4200
gttctcatg	taaattcagc	ttaacctggt	tatctataat	cttttattgg	caaagttaa	4260
ttctca	gtccataga	gatacagtgt	attttatgt	catacacaat	tagtctaatt	4320
cttgataatt	cagttaattt	agtttggcat	tttcctacca	cttactaaaa	ggtttacatt	4380
aatgactga	tttaatata	taggtgcaat	gttctatgtt	tattttatt	gttatgacat	4440
ttaagttagct	aatataattg	accgggtcta	agttctcctg	tttatccata	aatatgggtac	4500
attatggca	gtgttaatac	agtttttgc	tcattgcct	gtactttacc	agcagaccac	4560
agttttggcc	tgcgttagcc	aaccctcaga	acaaaatcat	cattccttgt	atttatattt	4620
gtatctgaga	tagtaaaca	gatggctg	caggtaaca	tggcacctt	attttttt	4680
ttaataggt	aaacttctc	aaaagtagct	tgctttgtat	aagaactaag	ctatcgtat	4740
agatatact	atccttggag	tttatgtt	agacaagaat	tatttactaa	aataaaataat	4800
aaacaagata	atgcattata	caatttggc	atttctcg	tctcaagtgt	atgcacat	4860
gtaaatataa	actaaccaca	agataggtag	attgattcat	tccattttaa	tctccttgc	4920
taattcagta	cctccataat	tgttctatc	ttcttccac	tgttacaaa	ttaccagtta	4980
attaactcgt	gaaagaaaaa	ttcacatatc	agaataaaaa	taaatgtata	ctcacttat	5040
aaaaatcacc	actgtctc	ttccttaata	ctagcagtgg	aaatgtaaat	ggcttactct	5100
acaaattttt	gtgtggcaa	atacatagc	aaactgttgg	gagctgc	atgttacattc	5160
ctcccttctt	atcccccttt	tcttccctc	acttttgc	ataacatatt	cctgtaccca	5220
aagcattcta	ccacagtct	atttgactcc	cacttgcata	aacttcttta	aaaaattcca	5280
tgtttaacca	tatgaccctg	tttgcttact	catattctcc	ctccctctcc	ccttccttcc	5340
tctcttcc	agaagtcatt	tgccctgg	gaaatatttt	gtagggattt	tttattat	5400
tattttagct	gatgaacctc	aggacaacgt	ctacacacac	acacatacat	acacgcacac	5460
aaaatctcg	ctgttgcaga	gtgggcttgg	aatcagactt	ctgtgtccag	taaaaaactc	5520
ctgcactgaa	gtcattgt	tttgactgt	tacagactga	ttccagtgaa	tttgatctaa	5580
tttctttgt	tctaatgtat	gtgtctgc	accttgc	ttttaattt	ataagctcca	5640
agtagttgt	aatttttgt	caactttaaa	tgagtttcat	tcacttctt	tacttaatgt	5700
ttaagtata	gtaccaataa	tttcattaa	ctgtttctaa	gtggtttanc	taccattctg	5760
ccattttta	tttttattta	attttatttgc	tttgagcaca	ctgatcaacc	actgaactgc	5820
cttctccat	tgtcctgca	tgtatataagg	tttacatttt	tgttatatg	gtttcatag	5880
ttgggatttc	anagnctga	tacanatata	tttcagtttgc	ttctctgggg	gaatttcatt	5940
tgcacatctat	tttttancta	tctgtatata	tttgatataat	ataaaaaaga	tattttgctt	6000
ctatttggAAC	atttgatatac	tgcacat	atttctgtaa	acagctgcag	tcaaaaataa	6060
aacactgaaa	ttttcaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	6107

<210> 48
<211> 703
<212> DNA
<213> Homo sapiens

<400> 48

ccacgcgtcc	gcaggacatc	gttttctaca	tggtggtgt	gttcctgacc	ttcctcatgc	60
tcttcgtgg	cagggtcac	ctggcatgg	ctctgggta	cctgggctt	tatgtgttct	120
atgtggtcac	tgtgattctc	tgcacctgga	tctaccaacg	gcaacggaga	ggatctctgt	180
tctgccccat	gccagttact	ccagagatcc	tctcagactc	cgaggaggac	cgggtatctt	240
ctaataccaa	cagctatgac	tacggtgatg	agtacccggcc	gctgttcttc	taccaggaga	300
ccacgctca	gatcctggtc	cggggccctca	atcccggta	ttacatgaag	tggagaaggaa	360
aatcagcata	ctggaaaagcc	ctcaagggtt	tcaagctgcc	tgtggagttc	ctgctgctcc	420
tcacagtccc	cgtcggtggac	ccggacaagg	atgaccagaa	ctgaaaacgg	ccccctcaact	480
gtctgcacatct	ggttatcagc	ccccctgggt	tggtctgtac	cctgcagtcg	gggacacctatg	540
gtgtctatga	gataggcggc	ctcggtcccg	tctgggtgt	ggtgtgtatc	gcaggccacag	600
ccttgcttc	agtgacctt	tttgccacat	ctgacagcca	gccccccagg	cttcactggc	660
tctttgcttt	cctgggttt	ctgaccagcg	ccctgtggat	caa		703

<210> 49
<211> 639
<212> DNA
<213> Homo sapiens

<400> 49

ggcacgagca	ttcacaggtt	acaaatgtcg	ctgccaactg	tcctggccaa	atgactctgc	60
atcacaacc	tttccttgca	tgtggagggg	atggatttac	tcagtcac	tttgcattggct	120
gcatcacttc	tgccttatgt	gttctggaaag	ctttaaagaa	tttatatttag	tgccttatatc	180
tttattctct	acatgtgtat	tgggtttta	ttttcacaaat	ttttctgttat	tgattatttt	240
gttttctatt	ttgctaagaa	aaattactgg	aaaattgttc	ttcacttatt	atcattttc	300
atgtggagta	taaaaatcaat	tttgcatttt	tgatagttac	aaccatgtc	agaatggaaa	360
ttcctcacac	cttgcacacctt	ccctactttt	ctgaattgtct	atgactactc	cttggggag	420
aaaaagtgg	actaaaaaaaaa	taacaaacga	ctctctcaaa	aaaattacat	taaatcacaa	480
taacagttt	tatgcaaaaa	acttgattat	ccttatgaaa	attcaattc	tgaataaaaga	540
ataatcacat	tatcaaagcc	ccccaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	600
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	639

<210> 50
<211> 867
<212> DNA
<213> Homo sapiens

<400> 50

ggcacgagca	ggtaactgggt	gactgcctgg	ctgaggaaaa	gttaactaga	cacttggggaa	60
aaggagatcc	aaggagatcc	gaggcaaaat	gcctttgtat	gcttttcttc	ctatctcttt	120
ttctttctct	ccttcact	cttccttc	cttccttc	tcctttcttc	ttcttttttt	180
tttctctttt	ccccccaccc	tctgcctgac	tccttccttc	cctcccttc	cctcccttc	240
ccctccctcc	ctccccccct	tccttccttc	cttccttc	tccttccttc	tcctccctcc	300
tcctctctcc	tccttccttc	gccttccttc	cttcgtctcg	ccaacttgcc	agaaggagcc	360
caagaaaaag	cacccagatg	cttcgttcaa	cttccttagaa	ttcttccttt	ttttatgttc	420
agaaaagatg	gaaattcatt	tctgcataag	agaaaagaaaa	aatttggaaaga	cagggtgtaa	480
gtgaacagggc	ccattataag	aaagaaaacaa	aatctatat	tctgtctaca	aggaagcggag	540
agagagaaaa	agagagaaga	aagaagtccc	aggattctaa	tgtacccaaag	ggatctccctt	600
tttcttgaaa	tgttctgaaa	atttcaccaa	aagagcacag	gagaacatct	tggcttattc	660

atggcgatg atgtaaagaaa actgagagaaa atgaaagaaaa tgaagaattt ctgtgcaga	720
taatacacag ccttgaggaa agaaaaggctt ttaagattat agatataaaag gctattgctg	780
tattctggaa taaaagaaaag tctgtatgtca gggaaagggg aagtggaaa aactggaaaa	840
agaaaaaaaaa aaaaaaaaaa aaaaaaaaa	867

<210> 51
<211> 1569
<212> DNA
<213> *Homo sapiens*

<220>
<221> SITE
<222> (341)
<223> n equals a,t,g, or c

<210> 52
<211> 1196
<212> DNA
<213> *Homo sapiens*

<220>
<221> SITE
<222> (590)
<223> n equals a,t,q, or c

<400> 52
gattggttct gtttatgtga tagattactt ttatttgattt gtatgttcaa ccagcccttgc

atccttaggaa tgaagccgac ttggttgtgg tggataagct tttgatgtg ctgctgggtt	120
tggctgccta gtgtttatt agggatttt gcgtcaatat tcatcaggga tattggccctg	180
gaattttttt tttttgttat gtgtctgcga ggttttgta tcagggtat gctggcctca	240
taaaaaataagt tagggagggc tccctgtt tcttcattt ggaagaattt cagaaggaaat	300
ggtaccagat ccycyttgtc cctctggtag aatttggctg tgaatccatc tggtcckgag	360
ctttttttt gtggtaggc tattaattac tgccctaatt tcagaacttg ttattggct	420
attcaggat ttgacttctt cctggtttag tcttggagg ttgtatgtgt gcaggaattt	480
attcatttct tctagatttt ctcgtttatt ttgttagagg tgttatagc atyctctgat	540
ggtagttgtt attctgtggg atcagtggtg atctcccctt tataatttt attgtgtcta	600
tttgattttt ctctcttktc ttctttatata ttctyctaa tggctatgtt attttggtaa	660
tctytacaa aaacaggctt cttagattcat ggatgtttt aaaggattytt cgtgtctcta	720
tcccttcag ttcttccctg atcttagcta ttcttctgtc tctgttagct tttgaaatgg	780
tttgccttgc ctctcttagt tcttttaacc gtgatgtcca gtgtgtcaat ttcatgtt	840
tccagccttc tgatatggc atttaatgtc ataaaattcc ctcttaaacac tgcttttagct	900
gtgtcttaga gattctggta cggtgtctt ttgttctcat tggtttcaaa taacttcatt	960
atttctgcct taattttgtt atttacccag cagtcattca agagcaggtt gttcaatttc	1020
catgtatgg tgggttttgg agtgaatttc ttaatcttga gttctaatgg gattgcactg	1080
tggctcgaga gacggttaca atttccatc ttttgcattt gctgagaagt gttttacttc	1140
caattgtgtc tcgtgccgaa ttcgatataca agcttacgta taccgtcgac ctcgag	1196

<210> 53

<211> 945

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (295)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (875)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (914)

<223> n equals a,t,g, or c

<400> 53

gaatggtaaa atattaatgt ctttctcccc caggttcagg attatgacag ctatgtccat	60
tcacctcttc tgcacatgtc tgcctgtgg aagtctggc cagtcataa aggcaattaa	120
aagaaataaa atatcaaacg attggaaaga tggtaatgtg tcatcattca tagaaaacat	180
gattcataga tatacataca cgaatgtttt gaattcataa gtagatttcg ccagtgtctg	240
gatataaagt caatatacaa aaactatttt tatagacatg aaacacgca tgagnaaaaaa	300
aatttaacca ttttagtag catcaaaaaa ccccccatacc tagaaatgt aatttggat	360
actatggg atatgttgcgat ggatattttt cattttccat ttgggattat tataaagaaa	420
atagccctga acatggtaa tataatgttgcgat ttggtaatgt tagcattcat ttctgttgat	480
tacaaactca ggggtaaaat tggtagtcc taaggagct atagatgtat tcaacttcag	540
ctgatatggc taaaataattt tgcaaaaaag attgcataa gttatgtcc catcagcaat	600
atgagagttc ctgttttcc acattgtcag caacacttg tactgttact ctttttaatt	660
ttagccgatt tggctgttgg tggtaataa tctcattgtc gtggccaggc gtgggtctca	720
cgcctgtat cccagcactg tggaaagcca aggtggcccg atcacgaggc caggagatcc	780
agaccatctt ggctaaatcg atgaaaccctt gttgcgttgc gtcccaacta cttggggaggc	840
tgaggcagga gaatggcatg aactcgggag gcggngctt cagtgagcct ccagcctggg	900
caacagagtg agantcttc aaaaaaaaaaaa aaaaaaaaaac tcgag	945

<210> 54
 <211> 488
 <212> DNA
 <213> Homo sapiens

<400> 54
 ggcacgagga gagtagaggc tattcatgtatgtctataaa aaaaataaca ccaaggctgg 60
 gattacaggc atgagccact gcacactggcc agtttgccta ttttgttgg tgcctccccc 120
 catggagac ctcaggagg tatgcctgcc ccacagatgc cctggaagga cagcttgctg 180
 ctctactgtaa catgactaa gagctgggtg gagacaatga acggtgagc catcggtccc 240
 gatgtggagg gagaacagct caagaccacg gaacagcctg ctctcccgct tcctggctc 300
 cgtgcgtttt tgcataatca ggcttttga ccaatcggcc aggccgcgcta tgtaaaatttc 360
 tgacattttc aaagctgtct tttaataaaa ctttcagtg taaaaaataaaa aaaaaaaaaa 420
 aaaaaaaaaa 480
 aaaaaaaaaa 488

<210> 55
 <211> 2860
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (753)
 <223> n equals a,t,g, or c

<400> 55
 ggcacacagg gctggcaggc cccgggtggc tgggtgttag gcatgaacaa attgtaccgg 60
 gtatccccca ccccaactctg accaccagg ctccttggat ttcactctcc cctgacaggc 120
 agcccaaccca ggcctggatt tgccctgtc tcccccctttt gctttttccc catgactaat 180
 gggcaccagg tcttgctgtc cctgcttc acctctgcag tggcagcagg cccctggccc 240
 caggtgcattt cccgtcagtgg gggttggatg tgcccttc caggcctgcc ctctgtccaa 300
 gcccggagtg ggcttgggtgg gctccctgtt ggcccccaactt ggggtgccagg tgggtggccc 360
 ggttatttag ggggtgggtgt atcaactgtatggacaggctt cttggcccaactt gctggagagc 420
 tgtttttttc aggaagggttc tggagatgaa gacttttttgcgaaatttccacc acaactccat 480
 gctggggggc tgggtctctg ctctcagacg cggacacca gggaggatag ccaggctgcc 540
 ctgcctggggaa atttgcatttgc gccgtcaat tcaacccgc ccaacgtggg caggaggcca 600
 cagtgccctg ccaggaggcag agggtctgaag gtctgcaggaa ggaaggacccat atcctgggg 660
 ggggcacccctg ctgccccaccc tgcccccaactt gtcctgggggg gggcacacc tgggcatttgg 720
 ggagttccagg gtcgtggggcc acacaagaga ggnngggggag aggccctggac agtaggaaga 780
 tcttgcctccagg gtcctggat ccgcctactt ggggttgacc ttggacaaac ctctgcctt 840
 gccctcagtc tccccatcaa gtttttcca ttccaggaggg tttggccat cttcagccac 900
 octaccagcc ctgaaaaggat tgcactctt gtttctggaa agtgtgtgtt gtgttaggt 960
 ggcctacaggc cctgggtgtt gggagggaaag gatggagaga cagcacagtg acagagccca 1020
 gactgcaggc tggaggatgg gttccacttc cccgcgtctg tgcgtcttgg accagtgcct 1080
 ctgaacccctt gcaacttgggg cagttggat taacatcttt ccaagccaa ttcttggggc 1140
 atcaggggctt ccgggtctt gggaggtggc aggtcttcag atggagatg ccatgggggg 1200
 gggaggtggcc ttcctttgg agggatgttggaa agtggagaca ggagtggctt ggcgcagctg 1260
 ccgtgtttctt taggggtctgg gccccggggag cccatggggc ttgtgccttag aaaggctgg 1320
 ctctctactg gggctctatgt gtcagactt catgtctccc cagctccagc tctgtttctt 1380
 ataggctcaag cttccacaaat gccagaggcc cagggcttagc cccctccacg tccctccat 1440
 atctacagct gcccctttaa tgacagcggc attgagttttt ctgggtgggg ggggtcatgc 1500
 aggggtggagg cagctgcctg ccgcgggtac tcattgcctg gccaggcagg acacaggctg 1560
 gcggggactg agagtggggcc ccacgaaatc cattgtcagg ttaccaggat gaagaaccca 1620
 ggctgtgtcggtt ggagtgcagg ccggaaagaat tatggcact gcagcaggag 1680

ggcagcctgg	gcatttagct	cctgatgtca	tcgatttggg	tgaggggaca	ggaaagtca	1740
aggaaactgg	ccagtggctc	tcacgcagac	ttacagcagt	ggagtgggtgc	ctgattccctg	1800
gtacagctgc	tccactgag	tctccaggg	tctgtggttc	aggaccccct	gcaaccccct	1860
cccagaccc	tgtactgggt	ggaggagagg	acctagagga	aagggtctgg	gcagataaagc	1920
agctgaggg	ggccctgggt	ttagcttatac	agtcttctgg	gcccttctgc	cccaggaagg	1980
gcagcgagga	ccatgggttt	gccccctgtca	tcgttatcgt	cctggccatg	agcttgcagg	2040
actggggagg	ccggagtcag	ccaggcagac	ggcagcacag	catttgcctg	ttggcaggtg	2100
gccttgggt	cttcccaaag	gcaatcgctc	cacgcagaac	aaaactcaact	tttttgggg	2160
gtgaagcacc	ttggttcatt	tgttagttc	gttaattcca	gcagtcgtt	tctaaggaa	2220
acatggctgc	agccggctct	gcccctccca	ccctccacc	agggtccccag	ttttcccaag	2280
ggcccccata	cccaacccata	ttcagggcgtc	agcatctctg	caccccaaata	gcctgttagg	2340
gaggatagta	aggctgagc	cctccctggc	ccatcaaaag	ccagcagtga	gagaacaccc	2400
ccatctctct	gaggtgacct	tgttagggcag	tccgtctgt	ctggctggcc	tgggtgaggt	2460
gggcaggagc	caaggcctgg	cgccctgggg	tcgctggct	tgtctgcgt	gctgacttca	2520
tcctgatagt	accttgattt	tcctactgtg	acttcccctt	ctgtcgactt	cctcaccac	2580
ttaaaattc	cgtattgaga	gcagtttctt	aagttaaccc	aaatcctatt	cagaagaagg	2640
ttcttctgg	aagttgggag	ggcgaaaaac	aagtttagtc	acagaagact	actccatgtt	2700
tgagcttctg	tttcaaggaa	agtgagtaac	tgccggagga	gcctgtcccc	tctgcagtgt	2760
gtggtgttgc	cctgatactt	ttcagattga	ggtgttactt	acatgtataa	aatgcacag	2820
acttaagtgt	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	2860

<210> 56

<211> 1559

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1445)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1551)

<223> n equals a,t,g, or c

<400> 56

atcccagct	ggggagacag	cgtgctgggc	aggcgctgcc	gagaccttct	cctgcagctc	60
tacccatcgc	ggccggagct	gggggtgccc	gtgcctgagg	tcctactgtca	cacgcgaagg	120
gtgtccagca	gcacgcgtctg	caagctggac	ggactctatcc	accgcttcat	cacgctccctt	180
gcggacacca	gcgactcccg	ggcgttggag	aaccgagggg	cggtatccag	catggcctgc	240
cgaaagctgg	cgtgtggcgca	cccgctgtg	ctgctcaggc	acctgccccat	gatcgccggc	300
ctccctgcacg	gccgcaccca	cctcaacttc	caggagttcc	ggcagcagaa	ccacctgagc	360
tgcttcttc	acgtgctgggg	cctgctggag	ctgctgcage	cgcacgtt	ccgcagcag	420
caccagggg	cgtgtggga	ctgccttc	tccttcatcc	gcctgtgtct	gaattacagg	480
aagtccctcc	gcacatctggc	tcgccttc	aacaagtttgc	tgcagttcat	ccataagtac	540
attaccatca	atgccccacgc	agccatctcc	ttccctcaga	agcacgcgg	cccgctccac	600
gacctgtctt	tcgacaacag	tgacctgggt	atgctgaaat	ccctccctgc	agggctcagc	660
ctgcccagca	gggacgacag	gaccgaccca	ggcctggacg	aaggaggcga	ggaggagagc	720
tcagccggct	ccttgccttgc	ggtcagcg	tcccttca	ccctctgc	cgccggccgag	780
atggcccccct	acatgaaacgc	gttttccccc	ggccaaacgg	tggagatgt	gctggaggtt	840
ctgagtgaca	tagacgagat	gtcccgccgg	agacccgaga	tcctgagtt	tttctcgacc	900
aacctgcagc	ggctgtatgt	ctcgcccgag	gagtgttgc	gcaacctcgc	tttcagccctg	960
gccctgcgt	ccatgcagaa	cagccccagc	attgcagccg	cttccctgccc	cacgttcat	1020
tactgcctgg	gcagccagga	ctttgaggtt	gtgcagacgg	ccctccggaa	cctgcctgag	1080
tacgctctcc	tgtgccaaga	gcacgcggct	gtgctgtcc	accgggcctt	cctgggtggc	1140
atgtacggcc	agatggaccc	cagcgcgcag	atctccgagg	ccctgaggat	cctgcata	1200

gaggccgtga	tgtgagccctg	tggcagccga	ccccctcca	agccccggcc	cgtccccgtcc	1260
ccggggatcc	tcgaggcaaa	gcccgaggaaag	cgtggcggtt	gctggctgt	ccgaggaggt	1320
gagggcgcgg	agccctgagg	ccaggcaggc	ccaggagcaa	tactccgagc	cctgggggtgg	1380
ctccggcccg	gcgcgtggca	tcagggcccg	tccagaagaac	cctcattcac	cttctggggcc	1440
acagccctgc	gcccggccgc	ggatcccccc	ggcatggcc	tggctggtt	ttgaatgaaa	1500
cgaccatgtac	tgtcaaaaaaa	aaaaaaaaaa	aaaccgcrgg	ggggggcccg	nacccaatt	1559

<210> 57
<211> 2064
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (2001)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (2024)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (2049)
<223> n equals a,t,g, or c

<400> 57
atggggcgagg ctgcggggcc ccggcgcgca cgccccgcacc tctccccagc cctggcggtgg 60
gcccgccccg gcccaggcag caatggggtt cctgcagctg ctggcgtaar cggtgcgtgy 120
atccgaacac cgggtggctg gtgcagccga ggtcttcggg aatccagcg arggtcttat 180
tgaattttct gtggggaaat ttagatactt cgagctaat aggcccttc cagaggaagc 240
tattttgcat gatatttcaa gcaatgtgac ttttcttatt ttccaaatac actcacagta 300
tcagaataaca actgtttcct tttctccgag gcgttagatcc cccaccatgt gacgcgtgg 360
cagaccagga ctccaggtgg aggttgcagt atgatgtcta tcagtatttt ctgcctgaga 420
atgacctcac tgaggagatg ttgctgaagc atctgcagag gatggtcagt gtgcggcagg 480
tgaaggccag tgctctcaag gtggttaccc taacagctaa tgataagaca agtgtttct 540
tctccctccct cccggacaa ggtgtcatat acaatgtcat tttttggac cggcttctaa 600
atacatctgc tgccctacatt cctgcctcaca catacgttg cagtttggag gcaggagagg 660
gtagttgtgc ttcccttagga agagtgtttt cccaaagtgtt cttcaactttt tttgcctgc 720
ttgggttctt catttttttcc ttggacacca gattctggaa aacagaattt ttcttcata 780
gtcttatcat catggatttc ttcttttata tactgattac aagactgaca cctatcaagt 840
atgatgtgaa tctgattctg acagctgtca ctggaaacgt cggtggaaatg ttcttggtag 900
ctgtgtgggt gcgattttgga atcctctcga tctgcatttctt ctgtgtttgg 1060
ggttccatcat ctcgtcagtacttttttta ctccacttggg aaacctaaag atttttcatg 1020
atgatgtgtt attctgggtc actttcttgc gcatagttat cctcattcca gtatgtttca 1080
tgggtcgctt aagaataactg aacataactga cttgtggagt cattggctcc tattcggtgg 1140
tttttagccat tgacagttagc tggccacaa gccttctca catcaacttg aacgtactca 1200
agagagcgt caacaaggat ttccacagag ctttccaaatg tggccctttt caaaactaatg 1260
acttcattat cctggcgttta tggggcatgc tggctgttgc tggattacg ttacagattc 1320
gaagagagag aggacgaccg ttcttccctc cccaccatca caagttatgg aagcaagaga 1380
gagagcgcgg agtgacaaac attctggacc ctagctacca cattcctcca ttgagagaga 1440
ggctctatgg ccgattaacc cagattaag ggcttccca gaaggagcag ccagctggag 1500
agagaacgcc tttgtttctg tagatgcca ggggcttgggt cagtgtgcct cagctttgg 1560
gttcatgcct ggagtgggttca aacagtctt ggtcaagtc taataagaga tcaggcatat 1620
atatctgttc ttgcataat attatggtgc ctttatttgc atatggtaag ggtgtacttag 1680
gggatttagga tgattgtaa agaatgagaa agatgacca aagttgggtg gttagggaggc 1740

tttttcttat ttccaaatac ttgagaaaatt acctttgggt ttacaaatct atgatcaact	1800
tattcatta aatagataca taaaaaaaaat taaaactga ttcttctgca gagcactgg	1860
gtttttttt ataaccctt gaaaacaagtct tcacactga gcctgtctaa actttcgag	1920
ggagtttatt attgagtctt tatctgtac agtatttggta gatttaggaa tttgatactt	1980
aggcccttga attttagaat ncaaaaaaggg gagccagcca gacnnggggg ccaaccgga	2040
tccccaaacng ggaccagggg ggtc	2064

<210> 58
<211> 1050
<212> DNA
<213> Homo sapiens

<400> 58	
ccacgcgtcc gcccagccag tccgcggcgc cgagggccgg ctgcgtgggg cagcatggcg	60
gggtcgccgc tgctctgggg gccgcggggc gggggcgctcg gcctttgggt gctgctgctg	120
ctcgccctgt ttccggccgc ccccgccgc tgcgcggc cggtaaagga gccccgcggc	180
ctaaggcgcag cgtctccgc cttggctaga ctggcgcctc tcggcgccttc cggcggtcag	240
tgccccgagg tgaggcgccg ggggcgggtgc agacctggcg cggcgctgg cgcatctgt	300
ggagccgaac gtccaggagcg ggcgcggggc gagggcgaga ggctgaggat cgcaggcgc	360
gcgtctcgcc gcagctgtcg cgcgtctgg gcgcggggc caactctgtat ccggctctgg	420
gcctggacga cggccggcgc ggcgcgtcgc cgcagctgc tcgcgtctgc ctccgcgc	480
gccttgaccc tgccgcctc gaagcccaagc ttgtccccgc gcccgatcccc gccgcggcgc	540
tccgaccccg gccccggc tacgacgcg gccccgggg cccggatgtc gaggaggcag	600
gcgcacgac acccgacgtg gaccccgagc tggtaggtt ctgcgtgggg cggattcttg	660
cgggaagcgc ggactcccgag ggggtggcag ccccgccgc cctccggcgt gccgcgcacc	720
acgatgtggg ctctgagctg cccctgagg gcgtgtctgg ggcgtgtcg cgtgtaaac	780
gccttagagac cccggcgc caggtgcctg cacgcgcctt ctgcgcaccc tgagcactgc	840
cggatcccg tgacccctgg gacccagaag tgccccccgc atccgcac caggactgt	900
ccccggccac acgtcccgag caacttaccc cggccagcca gcctctcac ccgaggatcc	960
ctacccttg gcccccaat aaacatgtac tgaagcagca aaaaaaaaaa aaaaaaaaaa	1020
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1050

<210> 59
<211> 2533
<212> DNA
<213> Homo sapiens

<400> 59	
ccacgcgtcc gcctggcaac ccctaataatt tggatctct caatgctatt tgccttcatt	60
agcttgctcg ttatgcttcc cacttgggg attgtgtctt cctggctggt atggggagt	120
attctatattt tgcgtatgtt cataagactt ttgagattat ggaggacagc caaactacaa	180
gtgaccctaa aaaaatacag cgttcatttga agatggatcc ccacaaacag ccgagcttt	240
actaacctcg tggaaaaagc ttacgtctc attcaagaaa ccgaagtgtat ttccagagga	300
tttacactgg tcagtgtcg ttgccattt aataaagctg gacagcatcc aagtcatcgat	360
ctcatcggtc ttccggaaagc tgcgttcccgactctaaagag ccaacttcca agcagcaagg	420
ctagtcaccc tatatatgtt gaaaaactac cccctgaact ctgagatgtcaatgtaaacc	480
aactacatct gtgtgggtcc tttttaaagag ctggcccttg gacttagtga agagcagatt	540
tcagaaggagg aagcacataa ctttacatgtt ggcttcagcc tgccgtcatt aggggtttt	600
ttccaactctt ggggtggcaca gagttcagag ttcttcagac ggtagccctt attactttct	660
acagccaatt cacccttgg gcccatttactt actccacgcac ttctgcctca tcgtatctta	720
tctgtatgtca tcacaggctt acctcatgtt cattctgtctt gtttggaaaga gcttaagcgc	780
agctatgtatgtt cttatcggtt ctttggaaact cagcaccatgtt cagtaccgcgttgc	840
aaaactcaac agaagtcaag agaactgaat aatgttcaca cagcagtgcg tagcttgc	900
ctccatctgtca aagcattact gaatgaggtt ataatttttgaagatgtact tgaaaagctt	960
gtttgtacta aagaaacaca agaacttagtg tcagaggctt atcccatctt agaacagaaa	1020
ttaaagtgtca ttccggccca cgttcaagca agcaacaatt gctggaaaga ggccatttct	1080

caggtcgaca aactgctacg aagaatac a gataaaaaag gcaaggctga aatagcatgt	1140
gaaaaccac attgtacagt aagtacctt gaagcagcc actctacaca ttgcagacaa	1200
agatccatc ccagaggagc aggaattaga agcttatgta gatgatata atattatgt	1260
tgatttcaga aaggatgatt ttattactt gtctcaagaa gacaaagaga gacagaagcg	1320
tgagcatgaa gaatccaaga gggtgctcca agaattaaaa tctgtgctgg gatttaaagc	1380
tccagaggca gaaaggcaga agtggaaagca acttcttattt agtgatcatg tgtttcttca	1440
tatagctta aaattatgtt attgacatta tggaaagat ttatcaatga gagaaatgtg	1500
tctcttttc accgtgtttg aaatccttgt ctctgttaga cccagtggaa cccataagta	1560
attcagaacc atcaatgaat tcagatatgg gaaaagtca gaaaaatgt actgaagagg	1620
aaagtaataa atccgccccca acagacaatg aaataatgt gactgatgt ttatgtgaaa	1680
actctctaga agttaaaaat aaagataatt cttaaatgt agtcttcccc caaggagcag	1740
aagaaagaat gtgttaccaa tggagatgt aagatgaaacc acaagcagat ggaagtggtc	1800
tgaccatgc ccctccaact cccaggact cattacgccc ctccattaaag cagaggctgg	1860
cacggctaca gctgtcacca gattttactt tcactgttgg ccttgctgca gaagtggctg	1920
ctagatctct ctccttacc accatgcagg aacagactt tggatgtgag gaggaaagac	1980
aaataataga agaaaaataaa aatgagatag aagaaaaatgt agaaccaa gttcatatgaa	2040
gtgatattag attgttcctt ttacaaaatgt gtttagctt aagactggaa agggaaatatg	2100
agtgttaatgt tactatataaa aagctaaga tggatgtt cagaaagaaac cctggtttg	2160
ataactgtatc tggaaatgtt agttacactt aaatggcaga tcttttaggtt aaaaatggaga	2220
aaggtaaggc ctctttgaa taaaactgtt ttttattttgt ggccacaactt atcaatcttg	2280
gaaattctt aagtatTTT aataagaaatgtt gatattatcat ttcttgcac aatttgcac	2340
cttaagggtt tttggaaatgtt tctgttgcac aacattaaac atttagtattt actctttttt	2400
actgtattttc tgcgtttaat aactgcactt attatgtttaa taacaatgtt tttgtatTTT	2460
atttttgtttt ataccagtct taaagatcca ggttctgaat aaaaaatattt attgatacaaa	2520
aaaaaaaaaaa aaa	2533

<210> 60
<211> 899
<212> DNA
<213> Homo sapiens

<400> 60

ggcagatttc cgggcacccctt cgtgggcacc acagagcccg cctcccccacc cctgagcagc	60
acctcaccctt ccactgtgc ggccactatg cctgtggc cctctgtggc cagcctggcc	120
cctccgggg aggccctcgct ctgcctggaa gaggtggccc cccctgcccag tgggaccgc	180
aaagctcggtt tgctctatgtt ctacgaggca gcccacagca gttagctggc cctgctggct	240
gtatggctca tcaactgtcta cagccgtctt ggcacatggc ctgactgttgc catggcggag	300
agaggcaaca agaaggggcaaa ggtccctgtc acctacttgg aactgtctcg cttaggcggat	360
gccccccatcc ccccccgtt ctggccttagg caggagggca tggcgcact gccacttaac	420
ttgtttgtt tgtagcactt tggtttaggtt ggggagaaatgtt caccatccatc tggccctgccc	480
ccttagtcacc tagctgttag ggtgccttag gctgaatggc tccaccctcc cccagccctg	540
cttctgaccc ttggctctgg agccctgtcc ctgcctgtca tcccccggca ccccccctcc	600
caggctccac taaggagggaa ggggctgtct gcagcactgt cactcagcac ctggccagg	660
gtggggccgc cgccatgtggg ctcaggttgc cccagggttgc ctcagcgaga gcccgttcc	720
tcagttgttca aaagctgtcat cagggaaatgtt cggcaaggca cacagggttgc tggcagcccc	780
tggggactgg ggcgttgcctt tggggagggaa gggcttgcctt aggctgttgc tggggcccg	840
agcagcatct tccgggttca tccctccctcc ccaccctca cagctcaagc caagtccac	899

<210> 61
<211> 1079
<212> DNA
<213> Homo sapiens

<400> 61

tcgacccacg cgtccgggtt tcaccacgtt ggccaggctg gtctcaaaact cctgacttca	60
gttggcctcc caaagtgttgc ggattaaagg catgaggccac tggcccccggc tacccttct	120

aactctactt	ctagtttctt	gcttctggc	tgctgtata	ccaaacagga	atgttaact	180
ttctgtcagc	ttcaggcctt	tgcacatgca	gttcactttg	tctatcttgg	tttttattct	240
taggattttt	attctcctaa	gaagctttt	ctgaccagcc	taaaaacttac	gtaagccctg	300
ggtaggtgc	tatgcttatg	tcctccata	gcattttgca	tttgcattgt	ttgttaactct	360
taatgtacag	catcatgatt	gcctatttt	actttctgt	ttgttacagt	agactttat	420
ctctttaagg	acaggaactg	tgtcttgtt	agaatcccc	gagcttattt	agtacaatgg	480
ctatgcttat	aatttaagta	tttattgaac	aatgaaattt	ttcctaagcc	ctaaaacctt	540
gcaagatgtt	ttagtgcagg	aaactggcct	cggtgagtt	gaataactag	cacgagggtca	600
ctcacctaaa	aagtggtgag	gagggattaa	aatctaaatc	tgttagctg	taaagattgg	660
gctttttyc	ttgctgctgc	acatgactgc	yctctctcat	gttgcctgta	cacatccctg	720
tcaagtgttc	aaacagcccg	tgcctaacaa	ccccatccat	agcttctgag	gaaagttgtg	780
tcatcttgg	acagctctga	gagctgaac	gagttttgc	agaataattt	cccatctatt	840
ggtcttattt	tatgctttgg	agaatataac	tttatttcaa	aaaacaaatg	attcagaatt	900
tgtcatctcc	ttaagggtccg	tttattttgt	tatttcattc	cttcatttcac	tgataaccat	960
ttactgagca	ccagcctggg	caacatggt	agaacccatc	tctaccaatt	aaaaaaaaaa	1020
aaaaaaaggc	ggccgctcta	gaggatccaa	gcttacgtac	gcgtgcattgc	gacgtcata	1079

<210> 62

<211> 1928

<212> DNA

<213> Homo sapiens

<400> 62

ggcacagag	taggtctgcc	ggcgatggag	tggtgggcta	gctcgccgct	tcggctctgg	60
ctgctgttgt	tcctcctgcc	ctcagcgcag	ggccgcccaga	aggagtcagg	ttcaaaaatgg	120
aaagtattta	ttgaccaaaat	taacaggctt	ttggagaattt	acgaaccatg	ttcaagtcaa	180
aactgcagct	gctaccatgg	tgtcatagaa	gaggatctaa	ctcccttccg	aggaggcatc	240
tccaggaaga	tgtatggcaga	ggtagtcaga	cggaaagctag	ggacccacta	tcagatcact	300
aagaacagac	tgtaccggga	aaatgactgc	atgtttccct	caaggtgtag	tgggttttag	360
cacttattt	tggaaagtat	cgggcgtctc	cctgacatgg	agatgtgt	caatgtacga	420
gattatcctc	agtttccaa	atggatggag	cctgcatttc	cagtttctc	tttcagtaag	480
acatcagagt	accatgat	catgtatctt	gcttggacat	tttgggaaagg	gggacctgct	540
gtttggccaa	tttattcctac	aggcttggta	cgggtggacc	tcttcagaga	agatctggta	600
aggtcagcag	cacagtggcc	atggaaaaaaag	aaaaactcta	cagcatattt	ccgaggatca	660
aggacaagtc	cagaacgaga	tcctctcatt	tttctgtctc	ggaaaaaaccc	aaaacttgg	720
gatgcagaat	acaccaaaaaa	ccaggcctgg	aaatctatga	aagataccctt	aggaaagcca	780
gctgctaagg	atgtccatct	tgtggatcac	tgcaaaataca	agtatctgtt	taattttcga	840
ggcgtagctg	caagtttccg	gtttaaacac	ctcttcctgt	gtggctca	tgttttccat	900
gttgggtatg	atgggttaga	atttttttat	ccacagctga	agccatgggt	tcactatatc	960
ccagtcggaaa	cacatctctc	caatgtccaa	gagctgttac	aattttgtaaa	agcaaatgt	1020
gatgtagctc	aagagattgc	tgaaaagggg	agccagttt	tttaggaacc	tttgcagatg	1080
gatgacatca	cctgttaactg	ggagaaccc	tttgcatttt	actctaaattt	cctgtcttat	1140
aatgttaacga	gaaggaaaagg	ttatgtatcaa	attattccca	aatgtttgaa	aactgaacta	1200
tagtagtcat	cataggacca	tagtcttctt	tgtggcaaca	gatctcagat	atcctacggt	1260
gagaagctt	ccataagctt	ggcacctata	ccttgcataat	ctgctatcaa	gccaaataacc	1320
tggttttctt	tatcatgtctg	cacccagagc	aactcttgc	aaagattttaa	aatgtgtcta	1380
atacactgtat	atgaaggcgt	tcaacttttt	ggatgatata	ggacccagaaa	tcgtgagatg	1440
tggatttgt	acccaactct	acctttcattt	ttcttaagac	caatcacagc	tttgtgcctca	1500
gatcatccac	ctgtgtgag	ccatcactgt	gaaatgtact	gttgcattgt	gatgtgc	1560
tttgcatttt	tatggggagc	agaaaatttcg	tcattttggaa	gttagtacaa	tcattgtctgg	1620
aattgtgaaa	ttattcaagg	cgtgatctt	gtcactttat	tttattgtat	gaaacccat	1680
gggggtttatg	aaaaataactt	ggggatcatt	ctctgaatgg	tctttaggaag	cggtagccat	1740
gccatcaat	gatgttaggag	ttctttttt	taaaaccata	aactctgtt	ctcaggaggt	1800
ttctataatg	ccacatagaa	agaggccaaat	tgcattgtat	attattgcaat	ttggatttca	1860
ggttcccttt	ttgtgccttc	atgcctact	tcttaatgccc	tctctaaagc	aaaaaaaaaa	1920
aaaaaaaaaa						1928

<210> 63
<211> 781
<212> DNA
<213> Homo sapiens

<400> 63
ggcacgagat tttcagccctt tttggactgg tttctccaca ttttcgtgga ttttatctaac 60
tttggcttt gatgttggtg accttcagat tgggtctctg agtgaacatc ctttttgtg 120
atgttgatac tattcccttc tgtttgggtt tttgtttcc ttcttaacagt cagggccctc 180
tgctcgaggat ctgcgtggagt ttgggttggg tccactccag accctgtttg tctgggttt 240
gccagaggag gctgcagaat agcaatgatt gctgcgtgtt ttccctctgg aagctttgtc 300
ccagaggggc acccaccaga tgccagccag agctctcctg tatgaggtgt ctgttggccc 360
atacttggag gtgccttcca gtcaaggatc acaggatgtca ggtacccact tgaggaggca 420
ctctgtcccc tattcagagct cgaacactgt gctggagat ccactgttct ttccagagct 480
gtcagacagg gacgtttaag tctgcgtgaag ctatgcccac agctgcccct ttcccccagat 540
gctctgtccc agggagaagg gagttttatc tataagtctc tgactggggc tgctgcctt 600
tcttcagaga tgccctgccc caagacgggg actctagaga ggcaagtctgg ctgcagtggc 660
cttgcgtgaac tttgggtggc ttcacccagt tggacccatc ctgagccctt ttttttaccc 720
tgtgagggtt aaaaatgccta atcaaggctc agcaatggtg gatgccttc ccccccaccaa 780
g 781

<210> 64
<211> 1194
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1172)
<223> n equals a,t,g, or c

<400> 64
ggcacgagaa gacatggagt cttaaagtgtg atcagtggga gggggctggaa atcatttaga 60
ggcatcttca ttccaaaaaac caggagctga tactggctgt cagccaggac ttcaactgac 120
ctatgttagaa cctgtccatg tggcccttcc ttgcagtctc cccatgggg ctgggttggg 180
tttccatcaca gtccggcagc ttacttctaa gggcaagcat tccacgacaa cacagcagaa 240
gggcattggc ttttttacagt gaagtttggc aatctcatag cgctcgcttct gtcctactt 300
atttatttggt cagggcaatc acaaagatgt gcataaggctc aaagaaaaaga gacataaccc 360
cgaccaacgcg atgaaagaag tgacacggtc atgtttagtag aggagtgtgtt gggatgggg 420
atagggtctgtt gggcacctgc agaaaaacacgc atctgtata ggctgtcatg gaagcgcagg 480
atgggggattt agcctacctg aggggtcaat cagcaaaaggc ctctgggagg aagttagatc 540
ttcggtcttag gatgtgaagg gctaaaaggaa gaatgaggaa gagtttcagg gagaggaatc 600
aatgaaacgca gtccagagac gctgggtgat tggatggttt gcttcagttt gatgacaata 660
cagagggca aggagactgg tgcaggagaa gagagaaggt gccatgtgtct ctgggtcggt 720
tcttctatgc cagactccct tagaagagga gcagcttcca gtcaaggatc tcccaggaaac 780
acggaggctt gacaggacaa tggcagccaa tccctgtcc caaaactgggt acagtgggg 840
aaagctgcattt ggtcttagatc caccctgtcc cctggggccca gtatagaaga tcaaatttcaa 900
tctgccttccat cttatccaga taaagtaaaag gaagactggaa aaaaagaatc aatccacggc 960
tccatctgcc catgactttc tctgcgtatc cccgaggatc ctatggataa agagacggca 1020
cacggccatgt cccgacgctg tggaggtggg gagaccccgca aagtccacag gaaaagagg 1080
aagtgcgttc caccctggca tccgctattc tctgcgttcc tgcctcatcc tcaattcaga 1140
ccatgatggaa gctgattgtc tcccatatca tnccttggat tgaatggatcgag 1194

<210> 65
<211> 1677

<212> DNA
<213> Homo sapiens

```
<220>
<221> SITE
<222> (1012)
<223> n equals a,t,g, or c
```

<400> 65

ggtcgcgtgg	tgccatcaca	gttcactgca	gccttgcact	ccggggctca	agcaatccctc	60
ccacctcagc	cacttggta	gctgagacct	cagatatgtg	ccatcacacc	cagctgattt	120
ttaaaaatta	atttttgt	gagatagggt	ctcatatgtt	gccatgtctg	gtctcaaact	180
actgggttca	aatgtatcc	ctgcctcagc	cttccaaagt	actgggatta	caggcatgag	240
ccacatgcc	gggctggag	ggcgaatttt	gttcagtcta	aaagataaagt	ttttcatagc	300
tctggctgt	gtggggggga	gcagaggagt	gaatgattgt	cagttggag	ggtcagagt	360
gggtctgtc	cctagggtgr	agtragggt	ggcttaggt	asmcamcaca	gaggccctgt	420
tcagccccac	gtccccccc	tgtgctcccc	cctctctct	cctctctgc	aggcggtgra	480
ggtatcatca	ttcagcagat	ttcaccagag	gcagtgagg	aggcaggtac	ctgagccaga	540
attcagaatg	tcttatttctc	cacttgactc	tgccactaac	ttgttgtca	actttggccc	600
tttccccagg	ctttcatttt	cttttcttt	cttttcttt	ytttttttt	gaggcggagt	660
ctcgctatgt	tgcccaggct	ggagtgcagt	ggcgcagcat	catctcggt	cactgcaagc	720
tccacccct	gagttcacgc	cattctactg	cctcagcc	ccgagtagcc	gggactgcag	780
gcrcccacca	ccacggccgg	cttattttt	gtatttttag	tagagacagg	gtttcaccac	840
gttagccaaag	atggctcga	tctctgacc	tcgtgatcca	cccgctggg	cttcccaaag	900
tgtctggatt	acaaggcgt	gcaactgcgc	ccggccat	tcttaaatat	ctaataaaa	960
atataatagca	aatgcgtt	ttaaaactacg	aaataatgac	cacgcggaaa	antattatct	1020
tccaagactg	ctggtccaag	gaaaagtcag	taataaaagt	gaagcattgt	agcttatgga	1080
atgactgggt	asatttggga	gaagccttag	caataatcta	gaatctgcat	agataataca	1140
tctgaggatt	gggctttgt	gttacaaaag	cattttttt	tcctctttt	atcccagccg	1200
tttgtctgga	ctgataaaaa	gcatttttat	tagttgtct	tattcaatcc	tcacaccacc	1260
tcaaattttac	agaggatatg	gatctggta	atttgtatga	ctatgtacc	tcatgtcagt	1320
ccacagca	gcctggaggt	ggtagaggt	ggtcctggc	tggaatccc	gccccagttg	1380
gaccttggc	aagttaactt	agctgtctgc	acctaattt	cctcaactggc	aaaacaggaa	1440
tactgggtt	tcacaccc	aattccagca	ctttgggggg	ctgagggtgg	aggattgttt	1500
gagtccagaa	gttcaaaaacc	agactggca	acatgcaag	accatctta	caaaaattaa	1560
ataaataaaaa	catttacaag	ggttgtgtg	aagattaaat	gagatcactc	acggaaaaagc	1620
tcagcagacc	ctgatgtca	gtaggtgtct	aataaatgtt	agccagcaaa	aaaaaaag	1677

<210> 66
<211> 1237
<212> DNA
<213> *Homo sapiens*

```

<400> 66
agcaaaaccca ggaagggtgtg ggtccccgc ttcgcgccaa gatggtgctg gtgctgcgc 60
atcctttgtg tgcccccggaa aggccgttcc gggagccggg tcgggggctc ctgactcgca 120
ctgggcagca tgacggtgcg cccgctgtca ctgctgtgcc gggacctctg ggcgtgtgg 180
ctgctgtctga aggccggcgc agtgcgtggg gcgcgggggg gtcctcgct ccccgaaagg 240
tgttgtgggg cgacatgcgg ggacccggg cgggggttgg a cgttctggc ccagccctgt 300
cctcagaagc tgctgggca gaagccggg gctggggat gccggggat ggttgggg 360
tgggtgcctc cgagaccga ggacccctgt tccttggcag ggaagggtgtg cacggccctt 420
gccccatgga tgggttaggg ccattggccct ggggtccctg gtgacgtg gggccgcctc 480
tggcccttggc ctgtgaggga ctgtctgtc tgggtcccaaga aggctgggat cacctttcca 540
ctggcccttcc tggtcgagggt ttttcatacga cagctatgt ggacaatga gggcagccgc 600
cacgtctggc tggtgaggg gctgcggctc ctcccttggag gggacgcctg gccactgtg 660
tccccacaaat gggccaccc gtggtcaag gctgtacaag ctgcctctc tagtaagca 720
ggacttggga ggcccttggc caagccgttgc gacccggctg ggcggccctct gtggtctcag 780

```

39

gtttgggtgt gtttgtctg	gtcagggctc	aggggctgct	ggtcacact	ggcccatcc	840
tgacaatttg	agctttgggg	caaggccc	ggagaagggg	tcacgtcg	900
ctgggttttgc	ttatgtctt	tctaagaatg	gagtactcg	tttcaagaga	960
ttatattttc	cagegggtac	ttatgccaag	tattgtgaa	taattcataa	1020
tttgtaat	ttatgtgaatc	agacctaacc	tatcaacggc	aatgaatgaa	1080
ttccaatttt	aaagtaaaga	actggctggg	tacagcgtt	cacgcctgt	1140
tttgggaggc	caaggctaga	ggatcgctt	agcccaggag	tttgagatca	1200
cataccaaga	cctcatctgt	aaaaaaaaa	aaaaaaaaa	aaaaaaa	1237

<210> 67
<211> 1934
<212> DNA
<213> Homo sapiens

<400> 67					
ccacgcgtcc	ggggcgttcc	tggcgttag	aggggagccc	caggggagct	60
gactgggtg	ataaatggcc	ggaaatttgg	cgtggccaca	ctcaacacca	120
ggagggcacac	tccgggtca	gcagcatcca	cagcagcatc	cgccatgtcc	180
ggggcctctg	atgegggtgc	tcgtggtcac	catgcctccc	atctactggg	240
agagagtggg	gaagccctga	atggccactc	tctgactggg	ggcaagtcc	300
acacgtggag	tttgcatacg	gggagctgtc	cacgatgacc	cagtggcccg	360
cccgatggcc	tctgtctt	cgacgtgggt	gtcaatgggg	ttgtccccgg	420
tgacgcagat	cttcaagtgc	aggacttta	ggagcactac	gtgcaaacag	480
gtgttcgtg	ggctccacac	agcgcttctt	ccagggccgc	ctccctctgt	540
caaccacagc	atccagtaca	acgcggcccg	gggccccccag	ccccagctgg	600
gcgggctca	gttatcagct	cggccttta	tccagaggcc	gaggccctgc	660
cgctacagcc	ctgcaggccg	aggagaacga	gtcggctgc	cccgagggt	720
ctcccaggga	gcttttgc	tggatgtgg	cgagtgtgc	tggatgtctc	780
agagggacag	cgctgtgtga	acctgtctgg	gtcctaccgc	tgcctccccc	840
tggcttccgg	gtggctgtg	ggggccgggt	tgaagatgt	gacgaatgcc	900
ggacgactgt	cactacaacc	agctgtgg	gaacacccca	gggggtcacc	960
ccccaggggt	tacggatgc	agggccctag	cctgcctgc	ctagatgtca	1020
gcagctgccc	aaggcctgcg	cctaccatgt	ccacaaccc	cagggcagct	1080
gtgccccccca	ggccagaccc	tccttcgcga	cggcaaggcc	tgcacccac	1140
tggacaaaat	gtgaccaccg	tcagccaccg	agggccctta	ttgcctggc	1200
ggcctcgatc	ccccgtaccc	cctaccacgc	ctgggtctct	ctccgtccgg	1260
cctgagcagt	gtggggccgg	cctggtcccc	tcctggttc	atcaggcaga	1320
cacagacccct	gacgagtgc	gctgtggaa	cacgcctgc	gcaacactga	1380
gggcagctac	cagtgcctgt	gccccccccc	ctaccgtct	ctcccccagcg	1440
ccaggacatc	aacgagtgc	aggaggagag	catcgagtgt	ggacccggcc	1500
caacacccgt	ggcagctacc	agtgtgtgg	cacaccctgt	cctgccaccc	1560
ccccagccct	gggacgtgt	tccggcgtc	ctcgaggac	tgcggcacgg	1620
tacgctgcag	tacccggctgc	tgccgctgc	cctggccgt	cgccccacc	1680
ccgcctcacc	gccttctccg	aggtcggcgt	ccccgcac	cgcacccgagc	1740
ggagcccgac	cccccgaccc	ccttcgcgt	gcttcgcgt	cgcggggcc	1800
ctacacccgt	cgccgcgtca	ccccgcggg	cctctaccgg	ctcaccgtgc	1860
accgcggccac	caaagcgtct	tcgtcttgc	catgcctgt	ccccctacc	1920
gggagaggc	attt				1934

<210> 68
<211> 3300
<212> DNA
<213> Homo sapiens

<220>
<221> SITE

<222> (1)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (3)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (15)
<223> n equals a,t,g, or c

<400> 68					
ncngcagccg gacgnccgag	cgcagcgagt	cagttagcga	ggaagcggaa	gagcgcggaa	60
tacgc当地 acc	gcctctccc	gcgcgttgc	cgattcatta	atgcagctgg	120
ttccccactg	gaaagcgggc	agttagcgc	acgcaattaa	tgttagtttg	180
aggcacccca	ggctttacac	tttatgcttc	cggtctgtat	gttgtgtgg	240
gataacaatt	tcacacagga	aacagctatg	accatgatta	cgccaagctc	300
ctcaactaaag	ggaacaaaag	ctggagctcc	accgcgggtgg	cgccgcgtct	360
gatccccccgg	gctgcaggaa	ttcggcaca	gaacacatct	taaggaaacc	420
agaaatcaag	taattatgaa	tgaacagatc	taaaaaagag	agagagaata	480
tcaacttagt	tgtgtttatg	accaaagaa	agatgttg	gtgttcaccc	540
agagatttc	ccttaaacct	cagttataa	tgaatggaa	tgaatgacag	600
tttcttcgt	ttcccaaaac	tctatcctt	tcttctcaaa	cagtggaaa	660
tacagtagac	agttagatgg	gagaacaaa	tcgtcatcca	agcatgttcc	720
ggtgttggag	agactctacg	cttccccc	ggatgttact	tcttctgtc	780
acctttgtt	cccttcattt	tgaggtgtgg	tcactcac	gtctaccact	840
ggcagctcg	gccttggtcc	catttggat	gatagatcac	atccctaata	900
tctgtgtcc	acactcccc	gctgactga	ccagttgttcc	cgccaaaacc	960
gacacttctc	cagggttttc	atttggtca	agcctactca	gactccaaac	1020
ttcagatcc	cggcagcgg	tgactgacat	gactgttgc	acggaaacgaa	1080
ggccaagagg	caaaatccat	gtttttgtac	cagactgtgt	tatattgtat	1140
attgacttgc	tgccctcaaca	gatctgcaaa	ggacaaccag	ccagttctgg	1200
cttctgggg	gaagtcagag	ggattatctc	aggatcagag	ctgataacgg	1260
ggccttcaag	gtgccaggcc	tgccccagta	cctccagagc	ctcaccagac	1320
tgca	ccggccggc	ccaaagatgg	agagcgggg	acgaatgtcc	1380
ctctcagctg	tttagaatctg	ccttccctga	agtcgcgtca	ctaacactgg	1440
ggaaaatctt	ttagcagcag	cctctggact	tggagagaag	ggcgtgcccac	1500
caacatggga	gagaaggatc	tatttggc	catgaaggaa	atcaccacc	1560
caagatctg	aaaattctcc	actgatgga	ccctgttg	tggttttttt	1620
ctgtgtccat	ctggccccc	aggatcttcc	gatctggac	agacggagca	1680
aagatctgaa	attcagatgt	tagtctgag	acttgc	tttccatgt	1740
gcagacatgt	gtggagaaca	ggaaattgtat	agctgctgag	cccaccat	1800
ggtcatctt	tcatgtgaag	accatcttcc	tacagatct	aggctggcc	1860
cctcaccat	actacaccac	tttcttctac	caacccccc	tcgttgaagt	1920
tacacttgc	ctctggaaat	gtgtcttac	ccttctgtc	agttgcagg	1980
agatgcagcc	acggaaaacc	tgacaaactgc	catgtcaca	aaatatact	2040
agatgttgc	ttctggccagg	tggatgc	catgc	ggccgtc	2100
gtgtgtatct	ctccagcagt	gggaccatgg	ggcccttgg	ttggccgtc	2160
gctgttggaa	gagagtgtat	acctcggtgc	ctgtgtgg	agatgtc	2220
agactacat	tttggaaaag	cagaagtca	cttttggcc	aggatgtca	2280
atacctctgc	aagcacctct	tctgtctc	ctcaaaatcc	ggctggcgtc	2340
tgagatgtc	tgtcaccc	aaaggatgtt	gtcagac	ccccaaagcc	2400
cttcagagag	cttccacc	ctgtgtgtt	tgtgaagaca	caagactac	2460
cattcaagag	gaaaggactt	tggcttgc	gtggagttca	2520	
agggaaagac	acccttagttc	gaggctgt	gcctttttgg	aaggaaagga	2580
tcttccaaga	acagaagccg	catgttgaag	aaaatctgg	gggggtatgt	2640

gtggatttt cctccactaa atctgcagga aacatgtga acataaattc	aaaaattta	2700
tccccaaaaa aaaaaaaaaa aaactcgagg gggggcccgg tacccaattc	gccctatagt	2760
gagtcttattt acaattcaact ggccgtcggt ttacaacgtc	gtgactggga	2820
gttacccaaac ttaatcgccct tgcaacat ccccccttcg	ccagctggcg	2880
gaggcccgca ccgatcgccc ttcccaacag ttgcgcagcc	tgaatggcga	2940
gtaaaggctta atatttgcgtt aaaatttcgg ttaaattttt	gttaaatcg ctcattttt	3000
aaccaatagg ccgaaaatcggtt caaaatccct tataaatcaa	aagaatagac	3060
ttgagtttgcgtt ttccagtttga	ccactattaa agaacgtgga	3120
aaaggcgaa aaaccgtcta tcagggcgat ggcccactac	gtgaaccatc	3180
agttttttgg ggtcgaggtt ccgtaaagca	ctaaatcgga	3240
tttagagctt gacggggaaa gccggcgaaac	accctaaagg	gagcccccgaa
gtggcgagaa	gtgtctcttt	3300

<210> 69
<211> 1797
<212> DNA
<213> Homo sapiens

<400> 69		
ggtcgacggt atcgataagc ttgatatacgaa attcctgca	cagttcttgg aaacccactc	60
gagaggccca cgccctccatt caccaggcaca	cgcatcacaa gaggcaacac	120
catgagctcg gggactgaac tgctgtggcc	cggagcagcg ctgctgtgc	180
ggcaggccagt ctgtgtgtgc	gctgctcacg cccaggtgca	240
ccagcagaga agtctgcgtg	aggaccaaca gagctttacg	300
ggtcggccag gcacatggccag	gacccttgcg	360
gttgcatttcc taccccttgc	tgaggatcc agcatcttcc	420
aggaagcaga cacgggttgcg	aggaaggctt catagaccccc	480
ctggggccgg ttctcgaagc	ccccagaaga tgatgtgcc	540
catttgcgaagc cagaaaaacca	aattcctacg agaatgtgc	600
cagaggggac ctcagcctgt	cactggccct gaagatggc	660
ctctgcctcc cccggaaag	atgaaggatcttgcgat	720
ccatttcgtg ggcgcgttcc	aggaaaggtaa	780
tggcccggtt ggaagccca	acggaggagga	840
ggcagccaca gaagccatgg	cggggaaaccg	900
tgtgcctatg gacccatcg	tgcccttccaa ggaccatttc	960
agcccttgc	ccagagctac tcaacttttta	1020
gcgtgcacag cctggaaaaa	gacagttact cacgggagct	1080
ctctccgcac ccaggctttt	gcaggccccg	1140
tatggacggc tgcgcaggat	tgccgttacc aagggttaacc	1200
gtcccaactt ggtatctgt	ttcttttgcgat	1260
tggtttgcgtt ggttttaaat	ctatgaattt ggattcttta	1320
atatttaaca ttctggattt	gattgtataa	1380
atgttagttttttccatc	tttgcgttccat	1440
acctccctt tctttggct	tcagttcaact	1500
tgtgcctcca gtggaaatgt	cacttgc	1560
agtgaccaga acaggcaga	agacaatagt	1620
ggctggccctc tgcagagctg	gttagtcccc	1680
cccagggga ctttgacta	ttcatggcc	1740
ctcaataaaaa ccatttttgt	tgaaaaaaaaaaaaa	1797

<210> 70
<211> 1373
<212> DNA
<213> Homo sapiens

<400> 70		
ggcacgaggg ctgacggcg	ctccatggcc	60
ttttgtctcc	gttgactt	
ggtgagttt	gtggcgaa	
gtttctgcgc		

42

tgttgcttag	taaccgactt	tcctccggac	tcctgcacga	cctgctccta	cagccggcga	120
tccactccc	gctgtcccc	cgagggtcca	gagggccttc	agaaggagaa	ggcagctcg	180
tttctctca	gaggagttag	gtcctttcag	ccatgaagca	tgtgttgaac	ctctacctgt	240
taggtgttgt	actgacccta	ctctccatct	tcgttagagt	gatggagtcc	ctagagggct	300
tactagagag	cccatcgcc	gggacccct	ggaccaccag	aagccaacta	gccaacacag	360
agcccaacaa	gggccttcca	gaccatccat	ccagaagcat	gtgataagac	ctcccttccat	420
actggccata	tttttggaa	ctgacccatga	catgtccaga	tgggagttccc	attctctagca	480
gacaagctga	gcacccttgt	aaccagagaa	ctattactag	gccttggaaaa	acctgtctaa	540
ctggatgtcc	atgtccctgg	caaggccctgt	ttagggccgtt	tgccgggttgc	catgctgtaa	600
atcctagcac	tttgggaggc	tgatgtgggt	ggatcacctg	aggtcaggag	ttaagaccag	660
cctcgccaaac	atggcgaaaac	cccatctcta	ctaaaaatac	aaaagtttgc	tgggtgtggt	720
ggcagaggcc	tgtaatccca	gctcccttggg	aggctgagggc	gggagaatttgc	tttgaaccgg	780
gggacggagg	ttgcagttag	ccgagatcgc	actgtctgtac	ccagccctggg	ccacagtgc	840
agactccatc	tcaaaaaaaaaa	aagaaaaagaa	aaaggctgtt	taatgcacag	gtgtgagtgg	900
attgtttatag	gctatgagat	aggttgatct	cgcccttacc	ccggggctcg	gtgtatgctg	960
tgctttccctc	agcagtatgg	ctctgcacatc	tcttagatgt	cccaacttca	gctgttggga	1020
gatgttgata	ttttcaaccc	tacttcccaa	acatctgtct	ggggttccct	tagtcttggaa	1080
tgtttatgtc	tcaatttttt	ggtgttggc	ctctcttcca	caagagttcc	tccatgttttgc	1140
gatagcgtt	gaagagtgt	tgggtgggt	gttgggatga	gatggagtgt	tcagtgc	1200
tttctcttatt	tacatttaa	agtcgttcc	ccaaacatagt	gtgtatttgtt	ctgaaggggg	1260
tggtggtatg	ccaaagccctg	ctcaagttat	ggacattgt	gccaccatgt	ggcttaaatg	1320
attttttctta	actaataaaag	tggatataat	atttaaaaaaa	aaaaaaaaaa	aaa	1373

<210> 71
<211> 1579
<212> DNA
<213> *Homo sapiens*

<400>	71					
ggcacgagga	tttggagggg	acaaaacatcc	aaaccattta	agtcacagca	ctttactccg	60
cagtgtaat	aacacaggca	tttcctata	taatcacagt	acagttatca	tactctggaa	120
attgaatatac	atctaataata	cttccatac	ccagatttc	ttagatttc	caatgatatt	180
tcttactgtc	ctccctttag	ccccccttctt	tctccattca	ggattctacc	attacatttc	240
atttcatgt	ctcttcagtc	tccttttagc	tttggtttttc	tttcttgatg	ttgccatctt	300
taggaggcca	ggccagggtt	tttgtaaag	atctgttctc	tttgatgt	ttcattttgg	360
atttgtttca	ttgttttgc	atgaatggat	tcaggctaaa	catttttggg	cagactgtt	420
tattgttata	cctagtgtat	tggtctttc	agtccatcat	ctggaggcac	ctgatggcag	480
ttttcccaat	attgcgaaat	taagtttgat	tattttgtta	aggttagtgc	caccagatct	540
ctccatTTA	aagacatct	tttctctaat	tactcagtgg	actgttaggt	gatgctttga	600
aactgaataaa	ctaacactcc	ctaactcagt	gatTTAGCAC	ccgttggattt	tttttttttg	660
cctgaatcaa	atattattat	agtagttta	aatggtgatt	ttccatttct	attttttgt	720
tagctgcccc	tcttcataa	ttttgtttt	atattttact	gggtgttaag	atgcatttc	780
cagctttct	ttgtttttta	cacctttcc	cattttttta	tttttttcca	tccttttttg	840
tcctgttttt	caatagatgg	atagaatttt	ctttctctgg	ttttaaagggtt	atacattttgt	900
gtgtgtgtgt	gtattctaaa	ccatTTGCC	ttaaaacata	gagatgtta	ttccgttga	960
ttaaaaaaaa	ctcagtaatg	ttactatctt	cttttaata	agataggtac	tttattttcat	1020
tctgtttct	ttgggtgaggt	ctcccttcac	ccagtcaagt	tgatgttaat	ctagaatttt	1080
tagttttta	aattatcgat	atactttctg	tcttttccct	tttttttccac	tctctctgt	1140
tttgtgtctt	ccccctttac	tctctttccc	ttcatttccct	ctttctcaact	ttcttcttct	1200
attttcaat	cttgggtaga	ttatctttta	gaaaacagca	agatattata	taatttactt	1260
tatattctt	ttccaaatga	ttaaaagtaat	aattaaaaat	ttttgatgt	tgtgtatgc	1320
aggatagggaa	tcctcttgc	atggaaagaa	ctctaccaca	tgcattggct	attagtgtgt	1380
taaacactgg	gaagtggctt	tagtccagc	tgggtgcct	gaagaaggta	ggtttcttca	1440
gttttcttat	tttaactgtct	ctttaccta	aaaaggagt	gaagaataact	gactcgacag	1500
gttttgcgt	gattccggta	acacagaagc	atagaactgg	aaaagaatat	taatatttgc	1560
caaaaaaaaaa	aaaaaaaaaa					1579

<210> 72	
<211> 1028	
<212> DNA	
<213> Homo sapiens	
<400> 72	
gcacgacaat tgaactgaac cctaaaaatg ctactcaat tcaccttatg ggtatttggt	60
gctatacatt tcccgaaatg ctttgtatc aaagaagaat tgctaaaatg ctgtttgcaa	120
ctgcctccat gtccaccta tgagaaggta gtatgttc ctttgttaag tttagtacgga	180
tttgttgaac cacagcgccc attctaccat gtgttccaca cattgtggag ctctggattc	240
agtgaagggg acttgaggca attcccttaa cgatccaatt caactgtgtt atcacaaggc	300
ttaacactta ttatccttga ctggtgagt gttttctttt tccccgttag gtgagtggt	360
ggtaattctg gaatactgtc atctaaaatg gctctggct aaaaatctacc ttcatttct	420
gtttgaaatc taaaactat tgaagtcat aaaaatagaaca agaaatacag catctgttac	480
ccagcatgtt tttagtgtat tacacacaat aacagaaaaag taaaagcagat gcttaagttg	540
ataaaaagaag aacactcatt ataacttcta ttttaaaaag catatgaaag gttcatattc	600
tctcatat tcaaggcctt ttgctttct tggtaaaaat aagatttgag aggaatttct	660
ggtttaaactt tggtttact catcacaagc ttttcagat aaaaaaacag gcaatcgaaa	720
aagctgtact ttttattttt acattataac aaggagcctt tttttcttc tggaaagcta	780
tagttagaa attgtatcaa aaaaactactt gttgtattct ttacacacag ttgagaata	840
tttattttat aatgcaccaa tttttatcaa tggttattttt aaaaatatgc ccatttgctg	900
gacacgggtt ctcatgcctg taatgccagc attttggaa gccaagggtt gtggatcagt	960
tgagcctggg agtttgagac cagcctggc aacgtggcaa aaccctgca aaaaaaaaaaa	1020
aaaaaaaaa	1028
<210> 73	
<211> 3674	
<212> DNA	
<213> Homo sapiens	
<400> 73	
ggcacgagct caaaagaaaat agggtgattt taaaaggatt aataaaaattc tgaaatgtta	60
agttagaagat tacattgtct agtcttgat ttcccttc ttttgtctc tttcattc	120
acactctcag tttctcatat ttgttagctca ttttatttggt tatttcctaa gaatattgaa	180
agtgaagcaa ctatgtgact gtattctca ggtaaaacact gactgcgtt gttggatttt	240
ccctat tttt gtgacttcaa gaataatgc ccctgtgaa tacatgcattt ttccatattc	300
gaaactgggtt agagtgggtt ggtgttctgc caacaattgc tagtgggtt aatttcatca	360
tatggccat tattgtctcatttccaaagaa ttcccttcattc aagcgttca gagctaggcc	420
agatcaatgc tacaatcatg aagttctcat tgcatcaat tggtaggt tgacaaggaa	480
ctcagataaa aatttccagg gtgcacttcc agaaccagct tcaacatatg tctacattgc	540
cccccaagttttaaataaagtgc aaccctttac tctctcatac agccagaaat gtttagaaatc	600
caaaaatcttgc gtgcattatt ttttcataaa cgctaaaaca ttgttggaaa caattttaatt	660
atttaaaattt caagtatattt attcacatata ttttcaatat ccaaatgttt aaaaattccc	720
agataattaa ctatgttattt cagatctcac cttaggggtt gatgttatga agactccagt	780
ggactgtact cacaatgttgc ctggacaccc tatgttggaaa ggttagaccc tcagcgaaaa	840
ataagaaggc cttttaccta cagggcagga cagggccca tgagagcagt tctgtggaga	900
tataaaaaggaa atggaaagaaatgcctta tagttagattt gatgttattt atctatataat	960
ctacatatat ctatctatct atatctatcatatctatataatc ttacattttt aattttagttc	1020
ctacacatata tggaaactct tctaataat tggaaatggaaa aatggaaaatggaaa aatataaaata	1080
ttccatggccccc aatgtggaaa tcaaaacatata taaaatttttggaaa aatggaaaattt tctttgttca	1140
cttctgttca aatgtggaaa tcaaaacatata taaaatttttggaaa aatggaaaattt tctttgttca	1200
tccatggtaa tacagataaa atacattttt ttaatttttggaaa aatggaaaattt tctttgttca	1260
gatcttccca gggaaaacat gtgttacatc tgcttacatc tccatggccccc tcttcttcc	1320
cctctccacg ttcccttgc taaatgtctg ggaaagcatg aacgttgcgtt caagaaccct	1380
gttgcgttccgg ctttttccctc ccctgtgaaa acgttactac tggtaggtt gatgtggaaa	1440
tgttagaaagg tggtaggttccgg ctttttccctc ccctgtgaaa acgttactac tggtaggtt gatgtggaaa	1500

cttgagaaaa aaacctaaga ttagccccagg tagttgcctg taacttcagt ttttctgcct	1560
gggtttgata tagtttaggg ttggggtag attaagatct aaattacate aggacaaaaga	1620
gacagactat taactccaca gtaattaaag gacgtatgtt ccatgtttat ttgttaaagc	1680
agtgtgaata gccttcaagc atgtgaataa tcttccatct tcccccac acatacacac	1740
acacacttt tgtttcttc aggttagacac ctttaaaat gcagaactaa ctgaggcatt	1800
tcaagtaacct tgctttcaaa tcaataaaagt caaatgtatg gaaacatttt gtgcctact	1860
ctccatacc cgtgtactca aattctctac tgtatgaatt atgcttaag tagaattcag	1920
tgccaaggag aacttggtga aataaaattat tttatatttt ttttattctt ttacaaaagcc	1980
atggatttta ttgggtttagt gtgtgcctcg tacacaagcc attcaatag gatggagctg	2040
ttaatttattt tccaaagagt aatagacatg caaaagttt aataaaaact gggccattaa	2100
caagtaaaattt aataaaactaa taagcattcc ctcttaggtt ttgcctaaac tgccttatcca	2160
ataacaaattt tgagaatctgt tgtaaaaaggt agttatattt cagagaaatg attttcattha	2220
ttgaaactgt tctcccttagc agggcatttt cccttttcc tgggatgtt gcaagtttag	2280
gagagaatag tcatgaaaag aaagggaga aaggggagaa gggaaagaggt taaaaagtaa	2340
gtgctcagac ctatgaacgt aatcccttt ctacaatata ttaagagcag ctgcgttgg	2400
ttgaaactga gttttgtcat ctccatatt tgcaggaaagg tattttctga ctgcataatgc	2460
agcttagatgt aaaattttat ttatcatcc tagaaagcct tgactagaaa aatgaataaaa	2520
tattgagggt ttccgttcca tatctggctt gcatgtgcca gaaagcagag aatagaaaat	2580
gtaatctcca acatccaagc atcgaaaccc aaggggttagg caattctatg taggttttg	2640
acatgaagtt tggtgcattc tgggttatgc tggctcaact gctattaaac ctctctggct	2700
tatagtcattc tcatttcattt agacaaggcac gtatcgaaca ctgcgttcgc acaaggctct	2760
ttagtttaca atttagcagc tactgtttgtt gttaaacaca cttttcacca aataggttct	2820
gaggcaaacg agagcaatga ctattnaaag aaaggcttcc ccagcatcac ttacacatcc	2880
caaaactaaa aagatcaact ctcccaactg agaaaagact cctggcttgc aatggaaaact	2940
tacagcagag agtcacaggc cacggcaaca acaacgacaa caacaaacat ttgaaatatt	3000
attctcaact cacgttttaa taatacatct tattttttt ctagtagaga aactacaaaat	3060
cagccttcc aacatttata tacagtttaa taaggcttgc gcaagttact ttttcttc	3120
cctgaggtat tttttcttc cccaccttgc ccctgttccct cccttcctct tctccctttg	3180
caagaggaaa tatttaacat atttgggtcc aacttcaata atgtataat taatacatta	3240
aaagcattta acttccttc tagaaaaatg cacaggctaa ggcatagaca aaacaaagag	3300
aaatgtcgag aaatttgcca ctggagacaa gcaatctgaa taaatatttg cccaaagttc	3360
tttttatgtc atatagtgtc aggattttgaa ggagcttattt ttttttaatg ttgcaactag	3420
caactcatct tcggaagaca cagccaggag aatgaagtag aagtggaaagg tttataaattc	3480
catttgcattt catttgcattt atatattttt aattcaagaa aaattgtgtt tatcttttaga	3540
attttgtatt caatacttta tgtactatgt gactcatgct tctggataaa taaagcaccac	3600
aatatgtatc tgaaccaca atcacacata ttatattaaa tatatatcta tataacaaaaa	3660
aaaaaaaaaaa aaaa	3674

<210> 74
<211> 2797
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (853)
<223> n equals a,t,g, or c

<400> 74

ggcacagagag agcagacaga attatatgtt gaggacacag gagatattt cattgtggat	60
ggagatggag gatttataa cagattgttc aaactgtccc aagatttcat gatcccttgg	120
ctgcattggggaaaatggggac agggcctgtt aagttcaaca tacctcacag ttttacactt	180
gattcagtcgt gtcgggtgtt ggttgcgtac cgaggaaata aaagaatcca agtattttgt	240
aaagacactg gggagggtt aggacatgg aataattgtt tcacagaaga ggaccttctt	300
cagtcattttt acttcctgtatg ggaagttactt gattgtggcc cagtcataatc ttagcaggct	360
ctcagtcgttta gcaacccccca cagtcgggaaat cattttttttt tttttttttt tttttttttt	420
ccaaacttagca gatcaagttt tgccacatct cttttttttt tttttttttt tttttttttt	480

45

tctatgttagc	agaaaattgg	gcaaaaacaag	tacaaaaata	tgtcccttgc	aatagctatg	540
ttccttcatt	tggttcataa	tgtttcttgc	ccgggaatat	ttcaagtggc	agttcagatt	600
ctcaattcac	taagtgc	aaaatgtgt	tcaagcacaa	gaatttattt	ttctagtata	660
aaagatctag	tatcagaaag	atttgtttt	gtatcattaa	gaatctata	tttgggtgcc	720
ctctgggac	ttagtttat	ttgtaagtgc	ataaggat	ttaatgaaa	ggaaaagtaac	780
aaaaaaaaatgg	gggttggaaag	aggactaag	gtggtaact	cattatttc	cctggtagac	840
tgattctccc	tgngtaaaaa	aaatggaaat	aaaatgac	ttgcatgata	atttattaaa	900
tttcatgtga	agaactccag	acctccagat	tgtgcacta	acataaaagt	agtcgttgc	960
gagatgtta	ataagatgaa	ctattgatta	atttgagtag	ccacagatg	ctgtgtcttg	1020
acgacttaaa	aatgaaaaag	catgattgcc	ttttagttag	tttgcagtct	agtggggaga	1080
caagcaggca	aacagtccaca	acacagcaaa	agcgaccc	gagcatagtg	ggacttttgg	1140
agtaggagtg	ctgcatttga	ctgagggaaat	catggatact	tcgcaggaga	agtgaaatttg	1200
agctcagact	tgaaaactga	ggaggagctt	accaagggac	aaggaggaga	aaacaataat	1260
ttccaagtaa	agaaggtata	aaaagttaga	agtgtactgt	aaactttgat	aggcttttag	1320
gccttttta	aagcccaact	tggctctgt	ccattaccta	taagatattt	aatgtcagtc	1380
agccctttaaa	tgttagga	aaatggctgg	catctaagca	ctttagtaaa	agaggttttt	1440
acaatataact	aaggattgt	gagtttctt	ctcttttttt	ttctttttct	ttctttttgtt	1500
ttatcatgaac	tcaacttatt	cctaaccat	gtctaccctta	aagaatattt	aagatttattt	1560
agataaactg	gatatgtgc	aaatccccc	agctgttaag	atgataattt	cctgtttcc	1620
tcctacatct	tctccccc	ctccctcc	ttgtgtgaat	attggcttcc	caattaagac	1680
ctttttttt	tccagtttgc	tttagttat	tatagtttt	ggaggaactt	tgccattttg	1740
taatcttca	aatcattctt	cacccttctt	cacatcagct	tcctgtttt	cccaagtgttt	1800
tactgtaaat	tgtgttagcat	atgacaaatc	ttgagctgac	tttcccttcc	acctgtttag	1860
gctggagttat	tttccagacc	tgaagggact	cacacttgtt	ttgatacttg	gatcacatct	1920
ccgtgaggtt	aggaaggtaa	atctaccac	aggaaggcct	gtactgtga	ttccaaggcc	1980
attggtaaat	gtgttgggtgc	cactgtatcg	actgtatgac	cttaaacaag	tcaccttagt	2040
tttcagtgtaa	atggaaatatt	cattgtctcc	tctttcatga	atgctgtgag	aatcagatgt	2100
gcaacaggta	catacttgc	cttggaaat	ctaataccctc	tgggatccaa	ttaagaggca	2160
tttaattaa	acaaaaggcc	ccttcttata	gtgttattta	tttgacataa	actatcatgat	2220
ttgccttaat	tttggttta	tagcatttat	ctaaacgtat	cctcatagac	tttatgcaga	2280
ttaatatgg	caatttgattt	ggataaaaaga	aagtaattt	agggtttgtt	tttaagccag	2340
gacaagaagt	gcaaattgc	cttggaaagca	atttaggc	aactgat	ttt gaaaatttca	2400
aatgttttat	tttactttgt	tttattaagc	caggacaaga	agtgcaa	atg ccttttggaa	2460
gcaattcagg	ctaggtaaac	cgattttgc	atttcaaaac	gttttatttt	actttgwttt	2520
rtrtcagagt	yttawaarvc	ctgtgtc	tatttctgaa	tgtctttgt	aaagtgttttgc	2580
ttagtgtacc	tgtgttattt	gtacttca	tttttctt	ggattaattt	gtttaatgaa	2640
tgagaaatgt	gttatgtttt	ttactaaaaa	gtataaatta	aaatttttgg	aagaaaaggc	2700
aatattatct	ggctccccaa	ttaaagtttgc	attttattgt	cacaaaaaaa	aaaaaaaaaaa	2760
aaaaaaaaaaa	aaaaaaaaaaa	aaaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaaa	2797

<210> 75

<211> 2703

<212> DNA

<213> Homo sapiens

<400> 75

ggcacgagat	ttcctacagg	tggaaacgccta	tcatttaggat	tcactgtAAC	gttagtgcta	60
ttaaactcAC	tagcatttt	attaatggcc	gttatctaca	ctaagtata	ctgcaacttg	120
aaaaaaAGagg	accttcaga	aaactcacaa	tctagcatga	ttaagcatgt	cgttggcta	180
atcttcacca	attgcatctt	tttctgcctt	gtggcgTTT	tttcatttgc	accattgatC	240
actgcaatCT	ctatcagccc	cggaaaataatg	aagtctgtta	ctctgtatTT	tttccatttg	300
cctgcttGCC	tgaatccagt	ctctgtatgtt	tttctcaacc	caaaggTTAA	agaagactgg	360
aagttaCTGA	agcgacgtgt	taccaagaaa	agtggatcag	tttcacgtttc	catcagtagc	420
caagggtggTT	gtctggaaaca	ggatttctac	tacgactgtg	gcatgtacatC	acatttgcag	480
ggcaaacctGA	ctgtttgcga	ctgtgcggaa	tcgtttctt	taacaaagcc	agtatcatgc	540
aaacacttGA	taaaatcaca	cagctgtcct	gcattggcag	tggcttcttG	ccaaagacct	600
gagggtctACT	ggtccgactg	tggcacacag	tcggccccact	ctgattatgc	agatgaagaa	660

gattcccttg	tctcagacag	ttctgaccag	gtgcaggcct	gtggacgagc	ctgcttctac	720
cagagttagag	gattcccttt	ggtgcgctat	gcttacaatc	taccaagagt	taaagactga	780
actactgtgt	gtgttaaccgt	ttcccccgtc	aaccaaaaatc	agtgtttata	gagtgaaccc	840
tattctcatc	tttcatctgg	gaagcacttc	tgtaatcaact	gcctgtgtc	acttagaaga	900
aggagaggtg	gcagtttatt	tctcaaacc	gtcattttca	aagaacagg	gcctaaat	960
taaatttgtg	aaaaatgcaa	tgtccaagca	atgtatgatc	tgtttggaaac	aatatatatga	1020
cttggaaaagg	atcttaggtg	tagtagagca	ataataatgtt	agttttttct	gatccatata	1080
aagcaaaatit	atacatttt	gtgttataag	cacaagat	agaacagctg	ttaatatttt	1140
ttaaaaatct	attttaaaat	gtgattttct	ataactgaag	aaaatatctt	gctaatttta	1200
cctaattgtt	catccttaat	ctcaggacaa	cttactgcag	ggccaaaaaa	gggactgtcc	1260
cagctagaac	tgtgagagta	tacataggca	ttactttatt	atgttttcac	ttgccatct	1320
tgacataaga	gaactataaa	ttttgtttaa	gcaatttata	aatctaaaac	ctgaagatgt	1380
ttttaaaaca	atattaacag	ctgttaggtt	aaaaaaaaatag	ctggacattt	gttttcagtc	1440
attatacatt	gcttttgtcc	aatcagtaat	ttttttcttaa	gtgtttgt	attacactac	1500
tagaaaaaaa	gtaaaaaggct	aattgtgtgt	tgggttttagt	cgattttgg	aaactactaa	1560
ctaatgtggg	ggtttataatag	tatctgggg	atttgggtgc	ttcatgtat	gttctcatta	1620
atgaatactt	cttaatatctcg	tttgcgttac	taataattttt	caatttgctg	ggatgtcacc	1680
tagcaatagc	ttggattata	tagaaagtaa	actgtggcata	ataactgcat	ttaatttagac	1740
aaaaacgggaa	gtatttatga	cacgaagttac	ttatgtttat	ttcttagtga	gctggattat	1800
cttgaacctg	tgcttataaa	tggaaatttc	catacatctt	ccccatacta	ttttttataa	1860
aagggcctat	tcaatagctc	agagggttggaa	ctctggtaa	acaagataat	atgttattaa	1920
taaaaataga	agaagaaaaga	ataaaagctt	gtcctgtgtc	ttttaaaatt	aaaatattta	1980
cttgattttcc	atctatgggc	tttagaccta	ttactgggtg	gagtcctaaa	gttataattt	2040
ttcaaatatgt	tttttgaaca	gtgtgtctaaa	tcaatagcaa	accctactgc	atattagtt	2100
ttctgtatatt	actaaaaaaa	tccagctaga	ttcaggtt	ataattaaac	tgtacataact	2160
gtgcataataa	tgaattttt	ttctttagttaa	atttttttta	gaacacaaagt	ttggaaatgt	2220
ggcttctgtt	catttcgttt	aattttaaagct	acccctctaaa	ctataagtggc	tgccagtagc	2280
agactgttaa	attgtgtgtt	atatactttt	tgcatgtttaa	atagtcttgc	ttgtacattt	2340
tcagtgtat	aaaaacagaa	tctttgtata	tcaaaatcat	gtagtttgc	taaaatgtgg	2400
gaaggattta	tttacagtgt	gttgtaattt	tgtaaaggcca	actatttaca	agttttaaaa	2460
attgtctatca	tgtatattta	cacatctgtat	aaatattttaa	tcataacttg	gtaaagaaact	2520
cctaattaaa	agggttttttc	caaatttcag	gttatttgc	tttttttattt	ttatttcattt	2580
aaaaactaga	ataacagata	tataaaatgt	ttaatcttgc	tgctatatgg	tatgaaatac	2640
aatattgtac	tcagtgttt	gaatttattaa	agttttctaga	aaggaaaaaa	aaaaaaaaaaa	2700
aaa						2703

<210> 76
<211> 742
<212> DNA
<213> *Homo sapiens*

<220>
<221> SITE
<222> (707)
<223> n equals a,t,q, or c

<220>
<221> SITE
<222> (724)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (726)
<223> n equals a,t,g, or c

<400> 76

gcgctcgaga atagtgggtc	ccccgggctg caggattcg	cacgagctca	cttcaatyc	60
tctttgagaa gttttccctt	tctccgcaac	cagatgtaca	tattgaact	120
ttggggggca ctctttcg	ggtagttctt	ttatTTTtat	taatctctgt	180
agtcctccaa caaccaaagg	ttgggactct	gtcttacata	tctgggtgcc	240
cagtaataag taagttgatt	atatacgagc	tatgttaactt	atattttta	300
atcaactgagt tttttttttt	aagaattttt	ttatttgaggt	aaacttcaca	360
ttaactatTT taaagtgaga	agttcagtg	cacttagtat	tgttaacaat	420
caccacccTT atttaaagtT	ccaaaaaaaaa	tgttctc	taaaaaggaaa	480
ttaagcagat actctccatt	ccttc	tccagcccc	agcaaccacc	540
ctgtctctat ggatttatct	attcttgcta	ttttatataa	atcgaattgt	600
tttgtctgg ctctttcac	tttagtacaag	tttttgagat	ttatTTcat	660
atcaacactt catttttag	gccaataaaa	attgttattat	gtgttntag	720
aananaaaaa atgaccctcg	ag			742

<210> 77
<211> 1825
<212> DNA
<213> Homo sapiens

<400> 77					
ggcacgagca tgcacatgt	atacctatgt	aacacacctg	cacattctgc	acatgtatcc	60
cagaacttaa attataataa	taaaaaaaga	ataattgggt	gatggcacat	ccagggttgc	120
caaagacagt cccagttat	gctgttgc	tggcattatt	aatatgaca	ctgccttaa	180
ctctcacaat taatttggat	gataacttat	atggtaactc	tgtctaaataa	aaaaaataaa	240
aattaccata gtaacaggaa	cctacttga	atgatgc	tgttctatt	ctggcttga	300
ttctgcattc ttgaggatt	tgtgcctca	tgacagaatc	ctatctacag	gtgatgtatt	360
tcatatgatt ttggctatt	tttttaacaa	tctcaagccc	aataatagcc	agtgtatataa	420
ggaatgtatgt tactttctcc	ccactttctg	gcaagttaa	tttagccacc	tgattacaag	480
aaggggacatt cagaggttagg	atggcacaaa	gacacagggt	ccactggaga	tcactggaa	540
cagctgcgc agggttaaga	gaaggggatc	ccagcagtc	ttcagtc	acacactaac	600
atcatcagtG aaaagtcc	gggcctgaag	atccagctat	gttgc	ttctca	660
ttaagtgaca gaacttggcc	caagcattga	ccatTTTGT	tcctcaataa	gcctgattca	720
accagggtca ccttgaatc	tgtcctccac	ctttccaata	aacctatTTT	atgcattcatt	780
cagttagtta ttatTTTT	tactttttt	ctgagaaaaca	tgactagatt	tagaaaaat	840
gtagaatttt acTTTTTTT	caatatttca	tgggtttcc	agagtttca	cgtgtttcac	900
accttccTTT gcttcccacc	atccccctt	ctatttggaa	ctagagagac	atgagtttga	960
attctagctg tgtaacctga	gtcagttatt	taaccttctt	ttgtttctgt	ttctttgtct	1020
gtaaatagca aaactacaa	ttaactttag	tccctgtgt	acaccaaaatg	ttatcttgc	1080
atattatata tattatATG	aattactact	gaaatgtct	aagatgcct	tgtgtgaatg	1140
gcattttgtt aaagattaaa	taatataagg	gaagtgtctg	ctttagtgc	tggcatataa	1200
taaaagctat tatttttacg	attatTTCC	atcttataga	agaattatcg	ttcttccctt	1260
ccaaagctaa taaatggaca	tgtgtttatc	agacagaac	taagagctc	caaataaaata	1320
ggaaataggt gctttcgaaa	gtctaggaaa	ataaaggta	ggaaattgtt	cataaaaattt	1380
agtacccata aatagcctat	aagttagatc	ccttagttat	tctatgcagg	aaaataaaagt	1440
tctacggagc acagattcca	aaactaattg	gtcataaaata	tcacctgaaa	gtttagaaaa	1500
tgtacgatca tggacctt	ttcatagtt	ctaaatctt	atatctgtgg	gatgggtcag	1560
gaatctagct ttgctaaagt	ccctcagatg	actctgtc	ttcttaggct	aaatacatgt	1620
ggtttggctt caatggacat	gttcc	gttgc	tgta	tcacacat	1680
tatgagagat gaggtc	tccat	ttcttaggtt	ctcttctc	cactccttac	1740
cctccatca ctacaataa	atcttttaga	aaattagct	tacatttgc	tcattataaa	1800
aaagaaagaa gataaaaaaa	aaaaaa				1825

<210> 78
<211> 1674
<212> DNA
<213> Homo sapiens

<400> 78

ggccacgaga gtagtgcgg cagctgcagg tcctggatt atttctcgat tcgctgtcgg	60
aggagaatga gaccctggtg gagtttgcta ttggaggccct gtgcacacctg tgcccagaca	120
ggccaaacaa ggagcacatc ctgcacgcag gaggtgtccc actcatcatc aactgcstat	180
ccagccccaa tgaggagacg gtgctgtctg ccatcaccac gctcatgcac ctgagccgc	240
cggccgcag ctttctcca gagctgaccc ccacgcctg ggtgcagtgc atgcttcgt	300
tctccctctc ggcgcagcgc aggttccgga acctggcaca gatcttcctg gaggacttc	360
gctccccccg ccaagtggcc gaggccgcga gccggcaggc gcaattttgc cctgggtatc	420
ccactgcgga ggagcgtggc cccacggcag cgctgatcca tggagactgc gagaccgtgg	480
caccctact gctggggacc acagtctcta tgtggacgc gggAACGGGG agcacatact	540
gccccattgg tgcccttca gccatctgaa aggccgggttc ttccagcagg acaggcatt	600
acactgatga aacgcccactg ggagtggaga agccagactc cagagacacg gagaagatca	660
aactggagct gcgttcatag gctggcactc tcaatctac atcaagggttgc accaccacca	720
gactcaggcc ctgggttaag aagccggccaa gtgcctggac ccagaggctt tgcaggacag	780
tgttctcagg agctggccct gaggcttagg agagctgcct tcgcctgcagg aaatcagggaa	840
ttatccctta acagaagtgt ctggagtagt ttccaggtat aggaatgaga tgcctcggt	900
tgaaaggatc tcacccttggg aagatgttgt gcccccttca gggctcttgg gatggatgc	960
ctccccccagg ggctctccaa gctggcatt tggcctggt ggatgccaac ctggataacc	1020
tgtggcccg cattgactgt ccacccagcc ttgcgtttag gcaccatgac tccaaagatg	1080
aagatgttgt ccctgcctt gagtgcacgc cccaggact taatgtggcc atcgggcac	1140
aagcacaagg ccatgcaggat gatgatacgt cggaatagag gcaccagccc tggtaactgc	1200
atcttccccc ctggccaccc catggccccc gctgaagct tcggccctcc tctgcgttca	1260
ctcaatgatg gggagcccta ccccagaagt gtatccacg agggcatcag ggacgcagtg	1320
agtgttgctc aaggggagtca ggaagagacg gcaacgtaaa ggatgtggct ccatgtccat	1380
ggtgccttgc ggtcaacata aggagcgtgg gatccgatgg aaaggtggag ctcagggaa-	1440
atgggggtcc ttgcctctcg tgcgttccctt caaggctgac cccttagatg gcccaggaaat	1500
ggcaggtgct aaaaaaatgg tacccacgtg ggcatggaaa tggggcagat taggggacca	1560
ctggacttag agggaggga agggctcatc agcaccgcct caggagcct gtcccttat	1620
gttcccaaataa aagggctt agaagactaa aaaaaaaaaaaaaaaa aaaaaaaa aaaa	1674

<210> 79

<211> 2191

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1327)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1334)

<223> n equals a,t,g, or c

<400> 79

ccttctctaa aaaagcaaaac aggcaaaact tcatgagaat ctgtatcatg taaaaatttt	60
atgtcttcg atttctccct acacacacac acacacacac acacacacac acacacacac	120
tcaacatttc ctccacccat atcatcactc cttagcatct ttatccatc aaaactttct	180
accccttgcatttctctgtc cagtttgcattt aatttccctc tcagcattct ctgttcaccc	240
ccacacccatg accctgaccc ctatcattt ctactacccca ggggtgttca cggttccagc	300
ctcctccatg aagcccaactt ctatggcctc actcccttgg gtaatgggg gccccccagct	360
atcatcttca ttgtatagaa aaccaactct gtgtatgtac ctgccttctc tccccttctc	420
tcctgaaaga gggctgggtt agagggtggg gggactggta tggcccttgc cgggtctgt	480
ttcgatctgg gaggagttt ggtactctgt gtagaaagaa atggggaggg gggaaatgggg	540
tggcccttgc acatccctaa gtccttgcct ttaatgttgc ctgttcgt	600

agcttcgaga ggagtggat	cgatcttctt gtggaaacgt	cctttcaat gtttacctgc	660
cgccaaacagg taggcactcc	caatggaaatg gaggggcgcg	gagggtggcc aaagactaca	720
tttccccataa ggctgcagct	ctcgggtgcc tgcgtgtgc	gtcctgagat acagtggaa	780
gtgttagttcc ctatcagatg	cttgggctga tgcttggaaa	ggaagttgga cacagcatt	840
cccatgaaac aatgggcaa	ctaactcttg aagctaaaaa	agatgtcctt ggaaccccat	900
ggggaaatttg ttatccccgg	tttgggttc ttttgttagg	gggggtttt gggaaaaactg	960
gggatttcctc cgatggaaag	ggggaaaaat attaaatagg	aagttattga cattaatgcc	1020
catgatagcc acccccactgg	gccatggaaag gtatgccc	gtgggtattt gaactaggct	1080
tttctgattt gtagaagtaa	cagagtaggg aaatttcatc	tacagcttta tttccctaaac	1140
tgcagtcagc acctgtaccc	tcatgaaagt tgccagat	aaagatctgt agtagtactt	1200
ttccaactta gttttccctt	gttttccca aaaacaat	ttttattttt ttttttttta	1260
tttaattttt tgagacaggg	tctggcttt tcacccaggc	tggagtgcag tggatgcac	1320
ttggctncac tgnacccctt	gectctcaga ttcagccat	ccttccaccc cagctctgcc	1380
actgagtagc tgagactaca	agcactcgcc accatgccc	gctaattaaa aaaataataa	1440
tcattttaaa tgcagcttt	attattataa tacaaaat	acatgaaaat aaaacccaaa	1500
catagcagtg ttattaaact	ctggcctgta gcagtgctc	acacctgtaa tcctagcagt	1560
ttggaggccg agacagggtgg	attacttgag acctggagtt	ttagaccagc ccaggtgaca	1620
cagcaagacc tcatctctac	taaaaataaa aaaaattag	ccaggtgtgg tggatgcac	1680
ctgtgtccca agctacttag	gatgctggag tgcgaggatc	gcttgagccc aggaggtaa	1740
ggctgcagtg aactatgtc	actcattaca ccccaagcctg	ggtagacagag cgagatgtcg	1800
tctcaaaaac aaacaaaacg	aaaaacaact ctggctagat	gctattgttt gccaagggtg	1860
cagtttccca ttattaaaaa	gtgaaaatata gggccaggc	cattggctca tgcctgttaat	1920
cccagcactt tggaggctg	aggtgggtgg atcacctgag	gtcaggagtt cgagaccagc	1980
ctggccaaca tggtaaacc	ttatctctgc caaaaatata	aaagattagc catgtgtcg	2040
ggtgggtgt tgtatctca	gctacttggg aggctgaggc	aggagaatca cttgaaccc	2100
ggaggcagag gttgcagtga	gccaagatg tgccattgca	ctccagcctg tgcaacgagc	2160
gaaactccaa ctcaaaaaaaaaa	aaaaaaa a		2191

<210> 80
<211> 1335
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1287)
<223> n equals a,t,g, or c

<400> 80			
ggatatatcc agggctgcgg	atttccccctt ctteagggtt	aatgtttctt gttttttctac	60
ctttccctcg cagtatacg	tcaacggcaa gawagtggaa	gttgcgtca aacagatcat	120
cgctgaaaaa gccgtggagc	aaggagggtgc tttctcgaac	cccgagaccc tggatctgta	180
ccgggacatc cctgagctgc	agggtttctg agtcagactg	gctggcgtgt cactcagccg	240
cacccgtgtg cactgttaact	tttgtgtgtc caagaaatta	tacagaaacc tacagcttt	300
gtaaaaggat gctgcacca	agtgttctgt aggcttgggg	agggategtt tctctgtttt	360
gttaaatctg gtgggttaccc	ggatcttcca caccgtgggg	attctggct tcagagacca	420
ggagggagtg tctggggccgc	agtgtggcac tgggtggaga	gtgtgtgtc ttgcacacac	480
agtgcagcgg gaacgggtgg	gctggctggt gctgaagaca	gacacactcc tgagccaagg	540
tcttgcttc aacctccccgg	tcccgttgc ccattttgtc	ctgtgaaggt gcaaattccct	600
ttcttccctt cccatctcg	gtcttcctgt tttccctcg	ggtcaggat gcctttgagc	660
tttagctgtt agaaaaggaaac	ccccgtgact tgacacagct	ttcacagct gctgcttagga	720
ccggccgggtt ggggtgttcc	gtgtgtctgt gtcatggat	caatgcaggc cctggaggac	780
tgtgcgtcac ccgtcaacca	gagcgtgcct ccggccagc	ttccctccaa ggaatggatg	840
gatttcatac aggtatcttt	tattgcacag actgaatggc	tttacatgtt tctaatgtga	900
attaggcatg tgaaggcagt	ggtgtccacc cgtgtccctc	atgggtgagc cctccagctg	960
tgagccagg cagtgtggtc	accgagttag gaccctctc	accaggaacc gcatccctgt	1020
gctgcctcca cctgagagtt	gctaggggt tcttgcag	atcatgtcat cagcacccct	1080

aagtcaagtc acgggttcc atagccaggc agttgtatg tacaatttag ttcagcgtat	1140
gaacttgtat ctctaatctg atgtccattt ttatatttt tgaaacttag cacaatgaaa	1200
tccttcttg aatcattttc cttttggatt ataaaaat gggggaaagt gctatgatga	1260
attttatgcataaaatgtat acatgtntgc acatgcaccc atgctaaaaaaa aaaaaaaaaa	1320
aaaaaaaaaaaaaaa	1335

<210> 81
<211> 1867
<212> DNA
<213> Homo sapiens

<400> 81	
cccacgcgtc cggcccacag cagagacagt ggagggcagt ggagaggacc gcgctgtcc	60
gctgtcacca agagctggag acaccatctc ccacccgagag tcacggcccc attggccct	120
cacctctcg tcctcgcccc catcctcc tc acgtggatg cttcccgagga ctggaaaggct	180
gaacgcagcc aagacccctt cgagaaaatgc atgcaggatc ctgactatga gcagctgctc	240
aagggtgtga cctgggggct caatcgacc ctgaagcccc agagggtat tttgggtggc	300
gctgggtgg cccggctgtt ggccgccaag gtgcctagcg atgctggaca caagggtcacc	360
atcctggagg cagataaacag gatcgggggc cgcacatctca cttaccggga ccagaacacg	420
ggctggattt gggagctgg agccatgcgc atgeccacgt ctcacaggat cttccacaag	480
ctctgccagg gcctggggct caacctgacc aagtccatccc agtacgacaa gaacacgtgg	540
acggagggtgc acgaagtggaa gtcgcgaac tatgtggggc agaagggtggc cgagaagctg	600
ggctacgcct tcgcgtcccc gggaaaggc cactcgcccc aagacatcta ccagatggct	660
ctcaaccagg cctcaaaaga cctcaaggca ctgggtcgca gaaaggcgat gaagaagttt	720
gaaaggcaca cgctcttggaa atatcttc ggggggggaa acctgagccg gccggccgtg	780
cagcttctgg gagacgtgtat gtcggaggat ggcttcttct atctcagctt cgccgaggcc	840
ctccggcccc acagctgcct cagcgacaga ctccagatca gcccgcacgt ggggtggctgg	900
gacctgtgc cgcgcgcgt gctgagctcg ctgtccggc ttgtgtctt gaaacgcgc	960
gtgggtggcga tgacccagg accgcacatgt gtcacgtgc agatcgagac ctctcccccg	1020
gcgcggaatc tgaagggtgtc gaaggccgac gtgggtgtgc tgacggcgag cggaccggcg	1080
gtgaagcgca tcacatttc gcccggctg ccccgccaca tgcaaggaggc gtcggggagg	1140
ctgcactacg tgccggccac caagggtgtc ctaaggttcc gcaaggccctt ctggcgcgag	1200
gagcacattt aaggcgccca ctcaaaccacc gatcgccgt cgccgcacatgt tttctacccg	1260
ccgcccgcgc agggcgccgt gctgctggcc tcgtacacgt ggtcgacgc ggcggcagcg	1320
ttcgcggct tgagccggga agaggcggtg cgcttggcgc tcgacgcacgt ggcggcattt	1380
cacggccctg tcgtgcgc gctctggac ggcacccggcg tcgtcaacgc ttggggcgag	1440
gaccacgaca gccagggtgg ctttggta cagccggcccg cgctctggca aaccgaaaag	1500
gatgactgga cggccctta tggccgcac tacatttgcg gcaagccacac cccctacccg	1560
cacggctggg tggagacggc ggtcaagttt ctgcggcccg ccatcaagat caacagccgg	1620
aaggggccgt catcgacac ggcacccccc gagggccacg catctgacat ggagggccag	1680
ggccatgtgc atgggggtgc cagcagcccc tcgcacatcc tggcaaaaggaa agaaggcagc	1740
caccctccat tccaaaggcca gttatcttc caaaacacga cccacacgc gacctcgcat	1800
taaagtattt tcggaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa	1860
aaaaaaaaaaa	1867

<210> 82
<211> 984
<212> DNA
<213> Homo sapiens

<400> 82	
gaattcgca cgagccacgc ggaagccaaag ccaccaggcc ccccaagcgac cacgcggagc	60
atgaacattt aggtatggcgc gtggccggc ctccccgtgc ccccccgtgc cgcccggtag	120
gatgtccctgg cccccacgggg cattgtctt cctctggctc ttctccccac ccctgggggc	180
cggtgaggt ggagtggccg tgacgtctgc cgccggaggg ggctccccgc cggccaccc	240
ctgcggccgt ggcctgtctt gcagcaacca ggcacccggc gtatctgcac cacggagaga	300

mctggccgag	gtcccagcca	gcatecccggt	caacacgccc	tacctgaacc	tgcaagagaa	360
cggcatccag	gtgatccgga	cggacacgtt	caagcacctg	cggcacctgg	agattctgca	420
gctgagcaag	aacctgggtc	gcaagatcg	ggtgggcggc	tccaacgggc	tgcccagcct	480
caacacgctg	gagcttttg	acaaccggct	gaccacggtg	cccacgcagg	ccttcgagta	540
cctgtccaag	ctgcgggagc	tctggctcg	gaacaacccc	atcgagagca	tccccctcta	600
cgccttcaac	cgcgtgcct	cgctgcggcg	cctggacact	ggcgagctca	ageggctgga	660
atacatctcg	gaggcggcct	tcgargggct	ggtcaacact	cgctacactca	acctgggcat	720
gtgcaacctc	aaggacatcc	ccaactgacg	gccctgggtc	gcctggagga	gctggagctg	780
tcgggcaacc	ggctggacct	gatccgccc	ggctccttcc	agggtctcac	cagcctgcgc	840
aagctgtggc	tcatgcacgc	ccaggttagcc	accatcgagc	gcaacgcctt	cgacgaccc	900
aagtgcgtgg	aggagctcaa	cctgtccac	aacaacctga	tgtcgctgcc	ccacgaccc	960
ttcacgcccc	tgcacccgct	cgta				984

<210> 83
<211> 2664
<212> DNA
<213> Homo sapiens

<400> 83						
ggttgctggc	ccaggtgagc	gggcgcgctg	gtccagggtga	ggggggcgct	ccccgcgacg	60
gcgcgtgcgt	cccgaggccg	ttcacgtaaa	gacagcgaga	tcttgaggc	cagccggaa	120
ggaggcgtgg	atatggagct	ggctgctgcc	aaagtccgggg	cccgcgccgc	tgccctagcgc	180
gtcctgggga	ctctgtgggg	acgcgcggcc	cgcccgccgt	cggggacccg	tagagcccg	240
cgctgcgcgc	atggccctgc	tctcgcgccc	cgcgctcacc	cttcctgctcc	tcctcatggc	300
cgctgttgtc	aggtgcagg	agcaggccca	gaccacccgac	tggagagcca	ccctgaagac	360
catccggAAC	ggcggtcata	agatagacac	gtacctgaac	gccgccttgg	acctcctggg	420
aggcgaggac	ggtctctgcc	agtataaatg	catgacggat	ctaaggcctt	cccacgttat	480
ggttataaac	ctcccccacc	gaatggatgt	ggctctccac	tgtttgggt	tcatcttaac	540
attgttatcc	cttccctgac	aaagtgtgc	aaccaacacg	acagggtcta	tgaracctgt	600
ggcaaaagca	agaatgactg	tgtatggaaa	ttccaggatt	gcctctccaa	gatctggcg	660
gatgtacaga	aaacactagg	tactaactcg	catgttcagg	catgtgaaac	aacagtggag	720
ctctgttttgc	acagtgttat	acattttagt	tgtaaaccat	atctggacag	ccaaacgagcc	780
gcatgcaggt	gtcattatga	agaaaaaact	gatcttaaa	ggagatgcc	acagctagt	840
acagatgaag	atggaaagaa	ataaccttg	acaataact	aatgttttta	caacataaaa	900
ctgtcttatt	tttgcgttt	gattatttg	agacctttaa	ataattttata	tcttgatgtt	960
aaaacctcaa	agcaaaaaaa	gtgagggaga	tagtgagggg	agggcacgct	tgtcttctca	1020
ggtatcttcc	ccacgttgc	tcccttactt	agtatgc	atgtcttgc	caatatcaa	1080
aacaagtgt	tgttttagccg	agaattttga	aaagaggaat	ataataactca	attttcacaa	1140
ccacatttac	aaaaaaaaaa	gatcaaatat	aaaatttc	ataatgtctg	ttcaacat	1200
tcttattttgg	aaaatgggg	aattatcact	tacaatgtt	tgtttactat	gaaattttaa	1260
atacacattt	atgccttagaa	ggaacggact	tttttttct	attttaatta	cacataat	1320
gtaatataag	tacaacataa	tatgttgc	ctctgttagcc	cgttggat	atgagtaat	1380
cacatttcta	ttaggactac	ttmcaaggac	aaggttcca	tttttccagt	tgtaaaattg	1440
gaaccatcg	ctgataacct	cgtagggac	aaacccagg	tagctaagt	ttatgtata	1500
tgcctagaag	gtgtatgtaa	tgcgattcg	aagcatagcc	actcccattt	tatgagctac	1560
tcacatgaca	aatgtcatct	tttgctataa	ccctttccaa	gttagagaaa	agatggattt	1620
aatgagataa	atgaaaagat	atttamcta	atataatcaag	gcactatttgc	ctgttatgt	1680
ttgttattta	tttcccagca	cttggcttcc	attgttagtt	ttttaaagac	tgtaacctt	1740
tactaactgt	ggtcttacta	aaattttgtc	ttgatactgc	tttcaaaaa	gcctttaatt	1800
agagccaaaa	ggatggaaaa	ggcaagat	aatgcctt	tatagatctc	ttatttacat	1860
tgaaaattat	taccatatgt	ttagagaaa	tccaagaaaa	cttcaacagc	ttctgaagat	1920
gtctatgaat	gttggaaact	tttcaatctc	ttggaaatgc	cagtttatgtt	cctagaccgg	1980
tctttgtca	ctactgggtt	ttaaccttcc	cctagcctgg	gacctcaagc	catatataatc	2040
ctttgggtga	cccatggcca	aagttattaa	gatgaactga	cttcaaagt	cagagaagga	2100
cagcataggg	agaggcggtt	atttgcgtt	cattacaggt	agaacagggc	agaaggaaaa	2160
gtatgttctg	gagaaaggc	catgttccct	actttggaga	tatgtcatttgc	ccgggaacct	2220
agtatcttcc	aacttgcattt	ggtggcagct	gttccagtg	gacaaggc	atgtatgcct	2280

tgtggctaag tgagcaaact	gggtttcac	ttaaatgtt	gggaccctca	attgattctt	2340
tatttcaac ctttataaaa	ggtacagttt	tgttaaggccat	tattaataat	taatgcttat	2400
cggctggca cagtggctca	cacctataat	cccacactt	gggaggctga	ggcggttgga	2460
tcacttgagg tcagggat	ttt gagaccagct	ggccaacatg	gtgaaacagc	gtctctacta	2520
aaaataaaaa aatttgcgg	gcgtgggtgc	gcatgtttat	agtctcagct	actcaggaag	2580
ctgaggtagc agaatacatt	gaacccagga	ggtggaggtt	gcagtgagct	gagattgtc	2640
cactgcactg cagcctggct	cgag				2664

<210> 84
<211> 1328
<212> DNA
<213> Homo sapiens

<400> 84					
cccacgcgtc cgggccagtg	gagggtccgca	gagtttggc	gccaggcgag	acggcaggc	60
ttaaagtcc gggaaatcaaa	gatcaactcc	cactgaggac	aaatggacct	gtaattccgg	120
gtgtgacgag agaacagat	ttaccttct	gaattaaaaaa	wcwgactccc	tgcgacaagg	180
actgtgtact gcatgaatga	ggctgagata	gttgatgtt	ccttggaaat	cctgattttag	240
agccgcacaaac aggaaaaaggc	ctgcgaggcag	ccggccctgg	cgggggctga	taaaccagar	300
cactccctc cctgtcccg	gtgcctcac	acaagttctg	ggagcagcag	tgaggaagag	360
gacagtggga aacaggcact	grtccaggc	ctcagccctt	cccagaggcc	gggggggtcc	420
agctctgcct gtacgaggag	cccttggagg	gaggagaa	aggatgtct	gaaatacgtc	480
cgggagatct tttttagtca	gggcataaaac	tgtgcactga	actgtctgcc	gagagcagct	540
ggaggacagc ttagtccca	ctgggtgtgc	tggccggccc	gcctgtggga	atggggctct	600
ctgtgctcct acctttgtc	cttcttggc	ctggcagatt	cacctcaggc	cagaagcccc	660
tggacactcc gggccttggg	gctgccgttc	tgagtgtgcg	gaaggcagga	ctaaaaatga	720
gatcccattt gactccctct	gtatgtactg	tgcctctcc	tgcgtcttga	ggctctggag	780
tcccaattgt ctgtgttagt	cagtgaccag	gttccaggaa	aatatgtgtc	atgtgggtgt	840
ccaaacttgc ggaaccaaag	agacagtact	ttgcaagaa	aaggatact	gccaggtgca	900
ctggaatttc tacagtttag	tccgcatgt	ctctctgaa	ggaggaagcc	tgtttcaaaa	960
ataggttcca tcatgagtat	atcaatgagc	tcccaccttc	ccagccagcc	tagaaagcaa	1020
acgagctgcc cacagttctc	tgcctgtct	gggagggtga	ggccacagt	tatagactgg	1080
taagccagac aggccctc	ccgcaagctg	ctaccttgc	ttcacctgta	ccttggtccc	1140
cgggcagcta gctataaaagc	aagaggacaca	ggagcccaga	agagacactg	aggacaagag	1200
atcacaccag agtacatgtc	tctgcctctg	ttttcagtgt	ggctttggac	aggaatatata	1260
gaataaaatca ctgcccataca	ggtttccaa	tacacaagtg	ctagaaaata	cacacaattc	1320
cccaatga					1328

<210> 85
<211> 1328
<212> DNA
<213> Homo sapiens

<400> 85					
ggccgcggca ggaggatttc	tgcctttgac	tgcaactctt	gtcgtcttat	gtgggtgttg	60
aattgtatctg tctctgcagc	cagatccagg	ctccctggaa	aaccatgtcc	ggcagctact	120
ggtcatgcga ggcacacact	gctgcccag	aggagctgt	gttgaat	tctgtgaatg	180
ttgggaagag gaatgcccaga	gctgcggct	aaaaattacc	caaccaagag	aaatctgcag	240
gatggacttt ctggctct	tcttgttca	cctggcttc	gtctgtatgg	gtcttggttct	300
tatctgcgtc tgctcgaaaa	cccatagct	gaaaggctc	gccaggggag	gagcacagat	360
attttccgt ataattccag	aatgtttca	gagagccrtg	catggttgc	ttcattact	420
tttccatacg agaaaccaca	ccttcattgt	cctgcacctg	gtcttgcaag	ggatgggtta	480
tactgagtac acctgggaa	gtatgttgc	actgtcagga	gctggagttt	tccttgcatt	540
accttcttc	gccttatctg	ctgcttagtg	taaacctgtt	tttttcacc	600
gaaccaatcc tggcattata	acaaaagcaa	atgaattatt	atttcttcat	gtttatgtat	660
ttgatgaagt gatgtttcca	aagaacgtg	ggtgctctac	tttgatttta	aggaaaccag	720

ctcgatccaa gcactgcagt gtgtgtaact ggtgttgca ccgtttcgac	780
tttgggtgaa caactgcattt ggggccttggaa acatcaggta cttcctcatc tacgtcttga	840
ccttgacggc ctccggctgc accgtcgcca ttgtgagcac cactttctg gtccacttgg	900
tgggtatgtc agatttatac caggagacctt acatcgatga ctttggacac ctccatgtta	960
tggacacggt ctttcttatt cagtagctgtt tcctgactttt tccacggatt gtcttcatgc	1020
tgggtttgt cgtggttctg agcttcctcc tgggtggcta cctgttggttt gtccctgtatc	1080
tggccggccac caaccagact actaacaggtt ggtacagagg tgactggggcc tggtgccagc	1140
gttgccttctt tggccgttgc cttccgttgc cagagccccaa agtccacccgg aacatttact	1200
cccatgggct tcggagcaac cttcaagaga tcttttctacc tgcccttcca tgcgtatgaga	1260
ggaagaaaaca agaatgacaa gtgtatgact gccttggac tggtagttccc gtttatttac	1320
acatgtggat ctcgttttc ca	1342

<210> 86

<211> 1154

<212> DNA

<213> Homo sapiens

<400> 86

aagacaggaa aagctccagg ccgtggttct ccaaagtgtgg tccctggaca gcagcaacat	60
cacccaggag cctgttaggg aaggcacagc ctccggccct gccccagacc tgccagaatca	120
gaaactctgg ggttatactgc tgtaacagac ctcccgatgg gttctgtatc	180
cctcttagagc aggagaacca cttagtttgc gggttgcgtt tggttggcat ctggccattt	240
gtgttagttc agaggaatgg ctgaccccca tgcgttcattt ctaagcttca ggcagctttt	300
ctccctggca gctgttcattt tgcgttgggg aatccctgggg actgtggctc ctccctccctg	360
tccctgtgtc cttgtatctgg cagttctaccc ctttcatttc cccgtggagg ctccatgcct	420
agaggtggtc tcaaaacaga agaatggcaa agataattgt ctgcgtttt accctgaccc	480
cattccctta agagggtcac ttcttggccc attcatttaa aaaccaatgt catatgttgc	540
tgtatccacc tatcagacag tgccacgtcc aaaggcgggg ctccyacctc cctgggraaga	600
gagactgttgc ctgtctgtgc ttccctgttt ctccagttccc acgctccacac ggacccacgc	660
ccttggagac tccctcrtgtc tccctgggtc tctgggtgtc tcaagacactt ccacactcaa	720
cgaccaactgg tgctgcagaa gggccgggtc ttacatttcca attaacagac gcttttccca	780
tctaattgcctt cttgccttctt cctaacaccca cctccggagt gtttatgtct attctaagtgc	840
aatttcactg tgcgtaaaaaa ttacacccctg ttgttcccgac gatttggggag gccggggcg	900
gtgtatcatt tggcccccagg agtttggggc tagctgggc aggttgggtga aaccccgct	960
ctataaaagaa atttttaaaaaa ttatgtgggc atatgtggcac gtgcctgttag ttccatctac	1020
tggggaggct ggggtggggag gatcgcatga gcccgggagt ttggggctgc agtgagctgt	1080
gatcgccggca ctgcacttcca gtctggccaa cagagcaaga ccctgtctct taaaaaaaaaa	1140
aaaaaaaaact cgag	1154

<210> 87

<211> 1197

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (573)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1177)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1185)

<223> n equals a,t,g, or c

<400> 87

aagacaggaa aagctccagg ccgtggttct caaagtgtgg tccctggaca gcagcaacat	60
cacctaggag cctgttaggg aaggcacgcg ctcagggcct gccccagacc tgcaagaatca	120
gaaactctgg ggtgaggcct ggttatctgc tgtaacagac ctcccagtgg gttctgtatgc	180
cctctagagc aggagaacca ctagcttgc ggttcgatg ttttggcat ttggccattt	240
gtgttagttc agaggaatgg ctgacccca tgtctcattt ctaagcttca ggcagctttt	300
ctccctggca gctgtcattt tgttgagggg aatccctgggg actgtggctc ctccctccctg	360
tccgtgtc cttgtatctgg cagtctaccc ccttcattctc cccgtggagg ctccatgccw	420
agaggtggtc ttcaaacaaga agaatggcaa arataattgt ctctgtttt accctgaccc	480
cattcttta agagggtc acggggccctt attcattaa aaaccaatgt catatctg	540
tgattccacc tatcagacag tgccacgtcc aangcggggc ttcacacccctt ctgggaagag	600
agactgttgc tgctgtgtct tcctgtgttc tccagtcaccc cgctcccaacg gacccacgcc	660
cttggagact ccctcagtgt cccagggtt ctgggtgtttt cagagaccc cacactcaac	720
gaccacttgt gctgcagaag ggccgggtct tacattccaa ttaacagacg cttttcccat	780
ctaattgcctc ttgccttctc ctaacaccac ctggggagtg ttatgtctt ttctaaatgt	840
attttactgt gtaaaaaaaat tcacacctgt tatccacgca atttgggagg ccgaggcggg	900
tgtatcattt gggcccaagga gtttgagact agcctggca agatggtaa accccgtctc	960
tataaagaaa ttttttttttggctggc tagtggcgcg tgctgttagt tccatctgt	1020
ggggaggctg gggggggagg atcgcatgag cccggggatgg tggggctgcg tgtagctgt	1080
atcgcggcac tgcactccag tctggcaac agagcaaaac cctgtctctt aaaaaaaaaaa	1140
aaaactcgag gggggggcccg gtacccaatt cgccctnats agtgnagtcg tattaca	1197

<210> 88

<211> 910

<212> DNA

<213> Homo sapiens

<400> 88

ggcagagctg gccttcgact cgctatgtcc actaacaata tgcggacc acggaggccg	60
aacaaagtgc tgaggtgagg accccagctg cgtggcaccg gtttgggtt gttgggttgg	120
atcggggccc tggaaagcgc ctgtctatcc cggggggcagg acctgagcgc ccctgaccct	180
cgagcctgtc gcaggtacaa gcccccccg agcgaatgtt accggccctt ggacgaccgg	240
acgcggact acatgaaccc gctgggcatg atcttcagca tgcggccctt catgcttaag	300
ctgaagtgtt gtgtttgggt cgctgtctac tgctccttca tcagctttgc caactctcg	360
agctcggagg acacgaagca aatgtatgtt agcttcatgtt gagactgtcc ctacagaaca	420
agtgtactttt ggttttttttggggacc ccagcgttgc catccttagac tgacaccctt	480
ctcccttctt catgtgtcc atcttcggcc tgggtatgtc ctatctgcag aatccctcage	540
ccatgacgcc cccatggta taccagccca gaagggtcac attttggacc ctgtctatcc	600
actaggccctg ggctttggct gctaaacctg ctgccttcag ctgccttccctt ggacttcctt	660
gaatgaggcc gtctcggtgc ccccaagctgg atagagggaa cctggccctt tccctaggaa	720
caccctaggcc ttacccttcc tgcctccctt cccctgcctt ctgcgtgggg agatgtgtc	780
catgtttcttta ggggtattca ttgttttttctt cgttggaaacc tttttttttttaaaat	840
actctgaaaa aaaaaaaaaaaa aaaaaaaaaaaat tygrgggggg gccccggaaacc caattcscgg	900
gatagtgtact	910

<210> 89

<211> 1076

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1029)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1037)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1040)

<223> n equals a,t,g, or c

<400> 89

ggcacgaggg gaaagccatg ctcccaggac tccttccttg cagccttaaa tcggtctgt	60
cggaaaattc cgcccttag aaacccacgc ttgggtgtaa ctattatttgc ttcttcctga	120
cctacttcctt gtttatcaact tccgggttca tcattttggc attcgggtga tcggggttgg	180
actattgaag cccgcttca ggttcttttc cccattttc cttgaaagg aagacttctg	240
gcttcctcta aatctccgtt ctctggtaa ggggagtcca agcctctgtc atgaggaacg	300
gaaatgcgag ggcctcggtt gttactctaa aatccgcct cagcttgcac gccggaaagct	360
gcgatttcgt cagcggaaaga ggcgtgatct ggccttcgac tcgctatgtc cactaacaat	420
atgtcggacc cacggaggcc gaacaaatgg ctgaggtaca agccccccg gagcgaatgt	480
aaccggccct tggacgaccg gacgcccggac tacatgaacc tgcgtggcat gatcttcagc	540
atgtcggccct tcatgcctaa gctgaatgg tgcgtggcat gatcttcagc	600
atcagcttttgc ccaactctcg gagctcgag gacacaagc aatatgttag tagcttcatg	660
ctgtccatct ctgcctgtgt gatgtccatct ctgcagaatc ctcagcccat gacgccccca	720
tggtgataacc agcctagaag ggtcacatcc tggacccctgt ctatccacta ggcctgggct	780
ttggctgctta aaccctgtgc cttcagctgc catcctggac ttcccgtaaat gaggccgtct	840
cgggtcccccc agctggatag agggAACCTG gccccttcctt agggAACACC ctaggcttac	900
ccctccctgcc tcccttcccccc tgcctgtgc tgggggagat gctgtccatg tttcttaggg	960
tattcatttg ctttctcggtt gaaacctgtt gtaataaaag ttttctactc tgaaaaaaaaaaaa	1020
aaaaaaaaana raaaacncgn gggggggccc ggaacccaat tcscggata gtgagt	1076

<210> 90

<211> 1842

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (67)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (98)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (212)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1838)

<223> n equals a,t,g, or c

<400> 90

gcgaccgcgc	ccttcagcta	gtcgctcgc	tgcgtctgt	tccctgtgc	cggctgcgc	60
tggcttnggc	gttggccgc	ctggcggcgg	ctcgagcngc	ctgcgsagcc	ggtaccagca	120
gttgcagaat	gaagaagagt	ctggagaacc	tgaacaggct	gcaggtgatg	ctcccccacc	180
ttacagcgc	atttctgcag	agagcgcaca	tnatttgac	tacaaggatg	agtctgggtt	240
tccaaagccc	ccatcttaca	atgttagctac	aacactgccc	agttatgtatg	aagcggagag	300
gaccaaggct	gaagctacta	tccttttggt	tcctgggaga	gatgaggatg	ttgtgggtcg	360
ggatgttta	gtatgtctg	accagctgag	gataggaaat	gatgggattt	tcatgtttaac	420
ttttttcatg	gcattccct	ttaactggat	tggttttttc	ctgtctttt	gcctgaccac	480
ttcagctgca	ggaaggatgt	gggccatttc	aggattttgt	ctctctctaa	ttaaaatggat	540
cctgatgttc	aggttttcca	cctatttccc	tggatattt	gatggtcagt	actggctctg	600
gtgggtgttc	cttgggttag	gctttctct	gtttctcaga	ggatttatca	attatgcaaa	660
agttcggaaag	atgcccggaaa	cttctcaaa	tctccccagg	accagagttc	tcttttattta	720
ttaaagatgt	tttctggcaa	aggccttct	gcatttatga	attctctctc	aagaagcaag	780
agaacacctg	caggaagtga	atcaagatgc	agaacacaga	ggaataatca	cctgttttaa	840
aaaaataaag	tactgttcaa	aagatcatt	ctctctat	gttcccttagt	gtaaaaatttt	900
aatagttaat	gcagaattct	gtatcttgc	aatcattatg	ggtaatgtt	tgaaaaaaagct	960
cttgcataat	agtctgtat	gtattaataa	tgccttat	attgtttgt	gtcatttttaa	1020
gttagcatgag	ccatgtccct	gtatgtcggt	ggggggcagtc	ttgtttttat	catctccat	1080
ctcaaaatga	acttggaaatt	aaatattgt	agatatgtat	aatgtctggcc	attttttaaagg	1140
ggttttctca	aaagtttaaac	ttttgtttag	actgtgtttt	tgcacataat	ccatatttgc	1200
tgttcaagtt	aatctagaaa	tttattcaat	tctgtatgaa	caccctggaa	caaaaatcata	1260
gtgcaaaaat	acattaagg	tgtggtcaa	aataagtctt	taatttggtaa	ataataagca	1320
ttaatttttt	atagcctgta	ttcacaattc	tgcggtac	tattgtac	aagggattct	1380
aaaggtgttg	tcactgtata	aaacagaaaag	cactaggata	caaataatgac	ttaattacta	1440
aaatgttaatt	cttgacactc	tttctataat	tagcggttctt	caccccccac	ccacaccccca	1500
cccccccttat	tttcccttttg	tccctgggt	ataggccaa	agtctgggg	taaggagagg	1560
atttaggtat	taggacaaa	gaaagaagta	gttggaaact	tttgagatg	tcccttaacat	1620
actgtactac	ttgtttttac	aatgtgttag	cagaaaccag	tggttataaa	tgttagaatga	1680
tgtgttttct	gccccaaatgg	taattcatct	tggtttgc	tgttaaaact	gttaataacaa	1740
cagaacattt	ataaaatatct	cttgcgttagc	accttttaaa	aaaaaaaaaa	aaaaaaaaaa	1800
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaanaa	aa		1842

<210> 91

<211> 1963

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (335)

<223> n equals a,t,q, or s

<220>

<221> SITE

<222> (1959)

<223> n equals a, t, g, or c

<400> 91

ggatccctcgc

ggcggccatc

cagcggccgc gccccttcag cta

cagttgc.

ccttacagca	gcatttctgc	agagagcgcga	gcatnatttt	gactacaagg	atgagtctgg	360
gtttccaaag	cccccatctt	acaatgttagc	tacaacactg	cccagttatg	atgaagcgga	420
gaggaccaag	gctgaagcta	ctatcccttt	ggttccctggg	agagatgagg	attttgtggg	480

tcggatcat	tttgcgtat	ctgaccagct	gaggatagga	aatgtatggaa	ttttcatgtt	540
aacttttc	atggcattcc	tcttaactg	gattgggtt	ttccctgtctt	tttgccctgac	600
cacttcagct	gcaggaaaggt	atggggccat	ttcaggattt	ggtctctctc	taattaaaatg	660
gatccgtt	gtcagggttt	ccaccttattt	ccctggatat	tttgcgtatggc	agtactggct	720
ctgggtgg	ttccctgttt	taggcttctt	cctgtttctc	agaggattta	tcaattatgc	780
aaaatcg	aagatgccag	aaacttttc	aaatctcccc	aggaccagag	ttctctttat	840
ttattaaaga	tgtttctgg	caaaggcctt	cctgcattta	tgaattctct	ctcaagaagc	900
aagagaacac	ctgcaggaag	tgaatcaaga	tgcagaacac	agaggaataa	tcacctgctt	960
taaaaaaaa	aagtactgtt	gaaaagatca	tttctctcta	tttgcgtatgg	tttgcgtatgg	1020
ttaatagtt	aatgcagaat	tctgtatca	ttgaatcatt	agtggttaat	ttttggaaaaaa	1080
gctcttgc	aaatggatgt	gatgtttaaa	taatgcctta	tatattgttt	gtatgtcattt	1140
taatgtat	gagccatgtc	cctgtatcg	gttagggggc	gtcttgcctt	attcatccctc	1200
catctcaaa	tgaacttgg	attaaatattt	gtaaatgtat	tatgtatgt	gccattttaa	1260
agggtttt	tcaaaaatgtt	aacttttgtt	atgactgtgt	tttgcgtatgg	aatccatattt	1320
tgctgttca	gttaatctag	aaattttttc	aattctgtat	gaacacccctgg	aagcaaaatc	1380
atagtgc	aaatacatttt	aggtgtggc	aaaaataagt	ctttaattgg	taaataataaa	1440
gcattaattt	tttatagcct	gtatttcacaa	ttctgcggta	ccttattgtt	cctaaggat	1500
tctaaagg	ttgtcactgt	ataaaacaga	aagcaactagg	atacaaatgt	agcttaat	1560
ctaaaatgt	attcttgaca	ctctttctat	aattagcgtt	cttcacccccc	accccccaccc	1620
ccacccccc	tattttcctt	ttgtctcc	gtgatgttgc	caaagtctgg	gagtaaggag	1680
aggat	acttaggagc	aaagaaaagaa	gtagcttgg	acttttgaga	tgatccctaa	1740
catactgtac	tatctgtt	tacaatgtgt	tagcagaaac	cagtgggtt	taatgttagaa	1800
tgatgtgtt	tctgcggca	tggtaatca	tcttgcgtt	ctatgtttaa	actgtaaata	1860
caacagaaca	ttaataaaa	tcttgcgtt	agcaccttta	aaaaaaaaaa	aaaaaaaaaa	1920
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1963

<210> 92
<211> 1487
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1487)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1487)
<223> n equals a,t,g, or c

<400> 92

gcgaccgcgc	ccctttcage	tagctcgctc	gctcgctctg	cttccctgt	gccggctgcg	60
catgcckwtg	gcgttggcgg	cgctggccgc	ggtcgagccg	gcctgcgcag	ccgggtaccag	120
cagttgcaga	atgaagaaga	gtctggagaa	cctgaacagg	ctgcaggat	tgctcctcca	180
ccttacagca	gcatttctgc	agagagcga	gttttccacc	tatccctgt	gatattttga	240
tggtcgtac	tgcgtctgg	gggtgttcc	tgttttaggc	tttctctgt	ttctcagagg	300
atttatcaat	tatgcaaaag	tccggaaagat	gccagaaact	tttcaaaatc	tccccaggac	360
cagatgttctc	tttattttt	aaagatgttt	tctggcaaag	gccttccgtc	atttatgtat	420
tctctctaa	gaagcaagag	aacacctgc	ggaagtgaat	caagatgc	aacacagagg	480
aataatcacc	tgctttaaaa	aaataaaatg	ctgttggaaa	gatcatttct	ctcttattgt	540
tccttaggtt	aaaattttaa	tagttaatgc	agaattctgt	aatcatttga	tcatttagtg	600
ttaatgttt	aaaaagctct	tgcaatcaag	tctgtatgt	attaataatg	ccttataata	660
tgtttgtat	cattttaaatgt	agcatgagcc	atgtccctgt	agtccgttag	gggcagtctt	720
gctttattca	tccctccatct	aaaaatgaac	ttggattaa	atattgtaa	atatgtataa	780
tgctggccat	tttaaagggg	ttttctcaaa	agttaaactt	ttgttatgac	tgtgttttg	840
cacataatcc	atatttgctg	ttcaagttaa	tctagaaatt	tattcaattc	tgtatgaaca	900

cctggaagca aaatcatagt gcaaaaatac attaaggtg tggtaaaaaa taagtctta	960
attgttaaat aataaggcatt aatttttat agcctgttatt cacaattctg cggtaccta	1020
ttgtaccaa gggattctaa aggtgttgc actgtataaa acagaaaagca ctaggataca	1080
aatgaagctt aattactaaa atgttaattt tgacacttctt tcataatta gcgttcttca	1140
ccccccccc cacccccacc ccccttattt tcctttgtc tcctgggtat taggccaag	1200
tctggagta aggagaggat taggtactta ggagcaaaga aagaagtagc ttggacttt	1260
tgagatgatc cctaacatac tggactactt gcttttacaa tggtagca gaaaccagtg	1320
ggttataatg tagaatgatg tgctttctgc ccaagtggta attcatctt gtttgctatg	1380
ttaaaactgt aaatacacaaca gaacattaat aaatatctct tggtagcac cttaaaaaaa	1440
aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaan cccggggggg ggccccn	1487

<210> 93

<211> 1653

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (67)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (212)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1636)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1653)

<223> n equals a,t,g, or c

<400> 93

gcgcaccgcgc ctttcagcta gtcgctcgcc tcgctctgtc tccctgctgc cggctgcgca	60
tggcttngc gttggccggc ctggccggcg ctgcagccgc ctgcgsagcc ggttaccagca	120
gttgcagaat gaagaagagt ctggagaacaa tgaacaggct gcaagggtatg ctccctccacc	180
ttacagcagc atttctgcag agagcgcaca tnattttgac tacaaggatg agtctgggtt	240
tccaaagccc ccatcttaca atgttagctac aacactgccc agttatgtt aagcggagag	300
gaccaaggct gaagctacta tccctttgt tcctggaga gatgaggatt ttgtgggtcg	360
ggatgatttt gatgatgctg accagctgag gatagaaat gatgggattt tcatgttaac	420
ttttttcatg gatcccttct ttaactggat tgggttttc ctgtctttt gcctgaccac	480
ttcagctgca ggaaggatgt gggccatttc aggatttggt ctctctctaa ttaaatggat	540
cctgattgtc aggttttcca cctatttccc tgcattttat aattctctt caagaagcaa	600
gagaacacctt gcagggaaatg aatcaagatg cagaacacag aggaataatc acctgttttta	660
aaaaaaaaaa gtactgttga aaagatcatt tctcttattt tggctctttagg tggaaaattt	720
taatagttaa tgcagaattc tggatcattt gatcattttag tggtaatgt tggaaaagc	780
tcttgcaatc aagtctgtga tggatcattt atgccttata tattttttt agtcatat	840
agtagcatga gccatgtccc tggatcggtt agggggcagt ctgtttttt tcatcctcca	900
tctcaaaatg aacttggaat taaaatgtt aagatatgtt taatgttggc cattttaaag	960
gggttttctc aaaatgtttaa cttttgttat gactgtgtt ttgcacataa tccatatttg	1020
ctgttcaagt taatcttagaa atttatttca ttctgtatga acacctggaa gcaaaatcat	1080
agtgcaaaaaa tacatttaaag gtgtggtcaa aaataagtct ttaattggta aataataagc	1140
attaattttt tatagcctgt attcacaattt ctgcgttacc ttattgtacc taagggatc	1200

taaagggttt gtcactgtat	aaaacagaaa	gcactaggat	acaaatgaag	cttaattact	1260
aaaatgtaat	tcttgacact	ctttctataa	ttagcgttct	tcaccccccac	1320
accccccta	ttttctttt	gtctcctgg	gattaggcca	aagtctggga	1380
gattaggtac	ttaggagcaa	agaaagaagt	agcttggAAC	ttttgagatg	1440
tactgtacta	cttgctttt	caatgttta	gcagaaaacca	gtgggttata	1500
atgtgcttc	tgcccaagtg	gtattcatc	ttgggttgc	atgttaaaac	1560
acagaacatt	aataaatatc	tcttgttag	caccccttaw	tgtaataaca	1620
aaaaaaaaaa	aaaaancccg	ggggggggcc	aaaaaaaaaa	aaaaaaaaaa	1653

<210> 94
<211> 1830
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (67)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (97)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (211)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1813)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1830)
<223> n equals a,t,g, or c

<400> 94
gcgaccgcgc cttcagcta gtcgctcgc tcgctctgtc tccctgctgc cggctgcgc 60
tggcttnggc gtggcgcg ctggcgccg tcgagcngcc tgcsagccg gtaccagcag 120
ttgcagaatg aagaagagtc tggagaacct gaacaggctg caggtgatgc tccctccac 180
tacagcagca ttctgcaga gagcgcacat natttgact acaaggatga gtctgggtt 240
ccaaagcccc catcttacaa ttagtctaca acactgccc gttatgtat aegggagagg 300
accaaggctg aagctactat ccctttgtt cctggagag atgaggattt tgtgggtcgg 360
gatgattttt atgatgctga ccagctgagg ataggaaatg atgggatttt catgttaact 420
ttttcatgg cattccttta taactggat ggggtttcc tgccttttg cctgaccact 480
tcagctgcag gaaggatgg ggccattca ggatttggc tctctctataa taaatggatc 540
ctgattgtca ggttttccac ctatccctt gatatttt atggtcagta ctggctctgg 600
tgggtgttcc ttgttttagg ctttctctgt tttctctagag gatttatcaa ttatgc 660
gttcggaaaga tgccagaaac ttctcaat ctccccaggaa ccaggttct ctatcttat 720
taaagatgtt ttctggcaaa ggccctctgt catttatgaa ttctctctca agaagcaaga 780
gaacacctgc aggaagtgaa tcaagatgca gaacacagag gaataatcac ctgctttaa 840
aaaataaaagt actgttggaaa agatcattc tctctatgg ttctcttaggtg taaaatttt 900
atagttaatg cagaattctg taatcattga atcatttagtg gtaatgttt gaaaaagctc 960
ttgcaatcaa gtcgtgtatg tattaataat gccttatata ttgtttgtatc tcatat 1020

tagcatgagc catgtccctg tagtcggtag ggggcagtct tgctttattc atcccccac	1080
tcaaaatgaa ctggaaatta aatattgtaa gatatgtata atgctggcca ttttaaagggg	1140
gttttctcaa aagttaaaact ttgttatga ctgtgtttt gcacataatc cataattgtct	1200
gttcaaggta atctagaaaat ttatcaatt ctgtatgaac acctggaagc aaaaatcatag	1260
tgcaaaaata catttaaggt gtggtcaaaa ataagtctt aatttgtaaa taataagcat	1320
taattttta tagcctgtat tcacaattct gcggtagcattt attgtaccta agggattcta	1380
aagggtgtgt cactgtataa aacagaagc actagagatac aaatgaagct taattactaa	1440
aatgttaattc ttgacactct ttctataatt agcgttctc acccccaccc ccacccccac	1500
cccccttatt ttcttttgc ttctgttgc tttaggcaaa gcttgggagt aaggagagga	1560
tttagtactt aggagcaaaag aaagaagtag ctggaaacctt ttgagatgat ccctaacata	1620
ctgtactact tgcctttaca atgtgttagc agaaaccagt gggttataat gttagaatgat	1680
gtgccttctg cccaaagtggt aattcatttt ggtttgtat gttaaaaactg taaatacaac	1740
agaacattaa taaaatctc ttgtgttagca cttttaaaaa aaaaaaaaaa aaaaaaaaaa	1800
aaaaaaaaaa aancgggggg gggggcccn	1830

<210> 95
<211> 1134
<212> DNA
<213> Homo sapiens

<400> 95

tccatctaca gtcctcacac aggtatttcag gaataccagg atggcgtgcc caagattcca	60
acagcctgta ttacgggtgaa agatgcagaa atgatgtcaa gaatggcttc tcatggatc	120
aaaattgtca ttccagctaaa gatggggca aagacctacc cagatactga ttcccttcaac	180
actgttagcag agatcactgg gagcaaatat ccagaaacagg ttgtactggt cagtggacat	240
ctggacagct gggatgttgg gcaggggtgcc atggatgtatgc ggggtggagc ctttatatca	300
tgggaagcac ttcacttat taaagatctt gggctcgctc caaagaggac tctggggctg	360
gtgctctgga ctgcagaaga acaagggtgga gttgggtgcct tccagtatttca tcaagttacac	420
aaggtaaata ttcccaacta cagtcgttgc atggagatctg acgcaggaaat ctttttaccc	480
actgggtctgc aatttcactgg cagtgaaaag gccaggcat catggaggag gttatgagcc	540
tgctcagcc cttcaatatac actcaggccc tgagccatgg agaaggagac gacatcaact	600
tttggatcca agtggagtg cctggagcca gtctacttgc tgacttatac aagtatttct	660
tctttcatca ctccccacggc gacaccatgc ctgtcatgg tccaaagcag atgaatgttgc	720
ctgctgtgt ttgggctgtt gtttcttgc ttgttgcaga catggaaagaa atgctgccta	780
ggtcctagaa acagtaagaa agaaacgttt tcatgttttgc ggcaggaaat cctgggtctg	840
caactttggaa aacttccttc tcaatataaca atttcatcca attcatcttc aaaggcacaac	900
tctatttcat gctttctgtt attatcttc ttgatacttt ccaaatttctc tgcattctag	960
aaaaaggaat cattctcccc tccctccac cacatagaat caacatatgg tagggattac	1020
agtggggca ttctttata tcacctttaaaa acatgttttgc ttccacttta aaagtaaaca	1080
cttaataaaat ttgtggaaaga tctctgaaaa aaaaaaaaaa aaaaaaaaaa aaaa	1134

<210> 96
<211> 1772
<212> DNA
<213> Homo sapiens

<400> 96

tcgaccacg cgtcggggag gatccccacg cgggtcccaa gcctgtgcct gagcctgagc	60
ctgagcctga gcccggccgg gagccggctcg cgggggctcc ggctgtggg accgctgggc	120
ccccacgtat ggcgaccctg tggggggggcc ttcttcggct tggtcttgc ctcagccctgt	180
cgtgcctggc gccttcgtg ctgtgtctgg cgcactgtca gacgcggccca aaaaatttgc	240
ggatgtcaga tggaaatgtt tctgtcccttc ctataaagaa aatttctggg catattttata	300
ataagaacat atctcagaaaa gattgtgatt gccttcatgt tggggagccc atgcctgtgc	360
ggggggctga tggaaagca tactgtctac gctgtgaatg caaatatgaa gaaagaagct	420
ctgtcacaat caagggttacc attataattt atcttcctat ttggggcctt ctacttctgt	480
acatggtata tcttactctg gttgagccca tactgttgcagag ggcgcctttt ggacatgcac	540

agtgtataca	gagtgtatgtat	gatattgggg	atcaccagcc	tttgcaaat	gcacacgatg	600
tgcttagcccg	ctcccgccagt	cgagccaacg	tgctgaacaa	ggtagaaatat	ggcacacgag	660
cgctggaga	tcagaatgtcc	agagcagcga	aaagtctgtc	tttgaccggc	atgttgcct	720
cagaatatt	gggaattgaa	ttcaagggtga	ctagaaaagaa	acaggcagac	aactggaaag	780
gaactgactg	ggtttgtctg	ggtttcattt	taataccctt	trgatttcac	caactgttgc	840
tggaaagattc	aaaactggaa	gkaaaaactt	gcttgatttt	tttttcttgt	taacgtataa	900
ata gagacat	ttttaaaagc	acacagctca	aagtccagcc	ataagtctt	tcctatTTT	960
gactttact	aataaaaata	aatctgcctg	taaaataaaat	taaaaaaatcc	tttacctgg	1020
acaagcactc	tcittttcac	cacatagttt	taacttgact	ttccaagata	attttcagg	1080
tttttgttgt	tggtgttttt	ttttgttgg	ttttgttggg	agaggggagg	gatgcctggg	1140
aagtggtaa	caactttttt	caagtcaact	tactaaacaa	actttgtaa	atagaccta	1200
ccttctattt	tcgagtttca	tttatatttt	gcagttgtac	cagccctcatc	aaagagctga	1260
cttactcatt	tgacttttgc	actgactgt	ttatctgggt	atctgtgt	tctgcactc	1320
atgtaaacg	ggatctaaa	tgccctgggg	cttttcacaa	aaagcagatt	ttcttcatgt	1380
actgtatgt	ctgatgcaat	gcattccatg	acaaactggc	catttgctag	tttactctaa	1440
agactaaaca	tagtcttgg	gtgtgtggc	ttactcatct	tctagttacct	ttaaggacaa	1500
atcctaagga	cttggacact	tgcaataaaag	aaattttatt	ttaaacccaa	gcctccctgg	1560
attgataata	tatacacatt	tgtcagcatt	tccgtcgtg	gtgagaggca	gctgtttgag	1620
ctccaatgt	tgcaagtttg	aacttagggct	gggggttgg	gtgcctcttc	tgaaaggct	1680
aaccatttt	ggataactgg	cttttttct	tcctctttgg	aatgtaaacaa	taaaaataat	1740
tttggaaaca	tcaaaaaaaaaa	aaaaaaaaaa	aa			1772

<210> 97
<211> 2381
<212> DNA
<213> Homo sapiens

<400> 97						
ccacgcgtcc	cgcaaggcca	gttcttagtgt	agagagaaaa	aggagccggc	agcggcttt	60
acgcgtcccc	gggctgcgc	ccactcttc	ggccgttaac	gcccgtctt	gcggctgtcg	120
tcaagcgcgg	cggtggccg	gcggggccgg	gctgaggggc	tgccatggcg	gcggccggcc	180
ggctcccgag	ctccctggcc	ctcttctcgc	cgctccctgc	aggcgttgc	ctactggag	240
tcggggccgt	cccagcgcgg	gcgcgtcaca	acgtcacggc	cgagctctt	ggggcccgagg	300
cctggggcac	ccttgcggct	ttcggggacc	tcaactccga	caagcagacg	gatctcttcg	360
tgctgcggga	aagaaatgac	ttaatctgt	ttttggcaga	ccagaatgca	ccctatTTA	420
aacccaaagt	aaaggatct	tcaagaatc	acagtgcatt	gataacaagt	gtagtcctg	480
gggattatga	tggagattc	caaattggat	tccttctgc	atatcttccc	aaaaattatg	540
ccaaaggtga	attaggagct	tttatctct	ggggacaaaa	tcaaaccat	gatcctaaca	600
atatgaccat	actcaatagg	acttttcaag	atggccact	aattatggat	ttcaatgg	660
atctaattcc	tgtatatttt	ggtacatcaa	atgaatccaa	ccagccacag	atactattag	720
gagggattt	atcatggcat	ccagcattga	ccactacaag	taaaatgcg	attccacatt	780
ctcatgcatt	tattgatctg	actgaagatt	ttacagcaga	tttattcctg	acgacattga	840
atgccaccac	tagtacctt	cagttgaaa	tatggaaaa	tttggatgga	aacttctcg	900
tcagtaactat	attggaaaaa	cctcaaaaata	tgtatgttgg	tggacagtca	gcatttgcag	960
actttgatgg	agatggacac	atggatcatt	tactgcccgg	ctgtgaagat	aaaaattggc	1020
aaaagagtac	catactactt	gtgagatctg	ggatgaagca	gtgggttcca	gtccttacaag	1080
atttcagcaa	taagggcaca	ctctggggct	ttgtgccc	ttgtggatgaa	cagcaaccaa	1140
ctggaaatacc	aatttcaatt	acccttcata	ttggagacta	caataatggat	ggctatccag	1200
acgctctgtt	catactaaag	aacacatctg	gaagcaacca	gcaggccctt	ttactggaga	1260
acgtcccttg	taataatgc	agctgtgaag	aggcgcgtc	aatgtttaaa	gtctactgg	1320
agctgacaga	cctaaatcaa	attaaggat	ccatgggttc	caccttctt	gacatttacg	1380
aagatggat	cttggacatt	gtatgtctaa	gtaaaggata	tacaaagaat	gattttgcca	1440
ttcatacact	aaaaaataac	tttgaagcag	atgcattttt	tgttaaagtt	attgttctt	1500
gtgtctgt	ttctaatgc	tgccctgt	gataacaccc	tttggagtga	atcaacctgg	1560
accttata	atgtatac	ctgttagatgc	aaatgggtat	ctgaaaaatg	gatcagctgg	1620
ccaaactcagc	caatccgcac	attagctct	ccaaactacca	tacaacgtgc	ttgggtttagg	1680
tcggagcgca	aattttctt	accatctct	cggttgcatt	ccccgtccat	ctggagaaaa	1740

62

atctatacga	aaacaagagt	ggactgcaat	cattccaaat	tcccagctaa	ttgtcattcc	1800
ataccctcac	aatgtccctc	gaagttggag	tgcccaaactg	tatcttacac	caagtaatat	1860
tgttctgctt	actgctatag	cctctatcg	tgtctgtgtt	ttcatcttgg	caataattgg	1920
cattttacat	tggcaggaaa	agaaaagcaga	tgatagagaa	aaacgacaag	aagccccaccg	1980
gtttcatttt	gatgcgtatgt	gacttgcctt	taatattaca	taatggaatg	gctgttcaact	2040
tgatttagttg	aaacacaaaat	tctggcttga	aaaaataggg	gagataaaat	attattttata	2100
aatgtatgtat	cccatgttga	ttatggaaa	gtattcaat	aaatatggtt	tgaatatgtc	2160
acaagggtctt	ttttttaaa	gcacttgttga	tataaaaaatt	tgggttctt	attctgttagt	2220
gctgttacatt	tttgttcctt	tgtggatgtt	gttgcatgtt	ctccagggtt	tgtgttattta	2280
taatcttatt	tgcatcatga	tgatggaaaa	agttgtgtaa	ataaaaataaa	ttaaatgagc	2340
agggaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	a		2381

<210> 98

<211> 1955

<212> DNA

<213> Homo sapiens

<400> 98

ggcgcggatgt	ccatccctgt	atttgctgcc	atgctttcc	ttttctccat	ggctacactg	60
ttagggacca	gcttcagtga	ccctggatgt	attccctcg	cgctaccaga	tgaaggcagct	120
ttcatagaaaa	tggagataga	agcttacaa	ggtgccgtc	cccaggcc	gcccccc	180
cctcgatata	agaatttcca	gataaaca	cagattgt	aactgaaata	ctgttacaca	240
tgcaagatct	tccggcctcc	ccgggctcc	cattgcagca	tctgtgacaa	ctgtgtggag	300
cgcttcgacc	atcaactgccc	ctgggtgggg	aatttgtgt	gaaaggaggaa	ctaccgc	360
ttctacccct	tcatccctt	tctctccct	ctcacaaatct	atgttccgc	cttcaacatc	420
gtctatgtgg	ccctcaaatc	tttggaaaatt	ggcttcttgg	agacattgaa	aggaaaatcc	480
tggaaactgtt	ctagaagttc	tcatatgtt	cttacactc	tggtccctcg	tgggactgtc	540
tggatttcat	acttttctcg	tggctctcaa	ccagacaacc	aatgaaagac	atcaaaggat	600
catggacacgg	gaagaatcgc	gtccagaatc	cctacagcca	tggcaatatt	gtgaagaact	660
gctgtgaagt	gctgtgtggc	cccttgc	ccagtgtgt	ggatcgaagg	ggtattttgc	720
caactggagga	aagtggaaat	cgacccccc	gtactcaaga	gaccagtagc	agcccttgc	780
cacagagccc	agccccccaca	gaacacctga	actcaaatga	gatgcccggag	gacagcagca	840
ctcccgaaaga	gatgcaccc	ccagagcccc	cagagccacc	acaggaggca	gctgaagctg	900
agaagtagcc	tatctatgg	agagactttt	gtttgtgtt	aattagggt	atgagagatt	960
tcagggatgt	agttaaaccc	gagacagaga	gcaagtaagc	tgtccctttt	aactgttttt	1020
ctttggctt	tagtcaccc	gttgcacact	ggcattttct	tgtgtcaagg	ttttttaaat	1080
ttctgtaaact	aaggcagtgg	cagaagatgt	cagtccaccc	tgataactgg	aaaaatgggt	1140
cttctggggc	ctggcactgg	tttccatgg	cctcagccac	agggtccccct	tggaccct	1200
cttttccctc	cagatcccg	cccttctgt	tggggtca	ggtctcatc	tggggctaaa	1260
agttttcgag	actggctaa	atccctccaa	gctgctgcac	gtgctgagtc	cagaggcagt	1320
cacagagacc	tctggccagg	ggatccta	tgggttctg	gggtcttcag	gactgaagag	1380
gagggagagt	ggggtcagaa	gatttccctg	gccaccaagt	gccagcatt	cccacaaatc	1440
ctttttagggaa	tgggacaggt	accttccact	agttgtattt	attagtgt	cttctccctt	1500
gtctcccatc	cactctgaca	ccttaagccc	cacttcc	ccattagata	tatgtaagta	1560
gttgttagtag	agataataat	tgacatttct	cgttagactac	ccagaaactt	tttaataacc	1620
tgtgcacatc	tcaataagaa	tttatgagat	gccagcggca	tagccctca	cactctctgt	1680
ctcatcttc	ctcccttctc	attagccccct	tttaatttgt	tttccctt	gactctctgt	1740
cccatttagga	gcaggatgg	cagtaaaaa	agtctgcact	tggtgcattt	ctttccctca	1800
gaggaaagct	gagtgcac	ttaaactaca	ttccctca	ctccctgt	gaggcctgca	1860
gaggccctga	atgcacaaat	ggggaaacca	ggcagcagaga	ggctctctc	tcctctcc	1920
tcccccgtat	taccctcaaa	aaaaaaaaaa	aaaaaa			1955

<210> 99

<211> 1958

<212> DNA

<213> Homo sapiens

<400> 99

ccacgcgtcc	ggggcgttcc	tggctgtgag	aggggagccc	caggggagct	ggggcagcat	60
gactggggtg	ataaatggcc	ggaaatttgg	cgtggccaca	ctcaacacca	gcgtgatgca	120
ggaggcacac	tccggggte	gcagcatcca	cagcgcata	cggcatgtcc	cagcaaact	180
ggggcctcg	atgcgggtgc	tcgtggtac	catgccttcc	atctactggg	ccctggccag	240
agagagtggg	gaagccctga	atggccactc	tctgactggg	ggcaagttcc	ggcagagtc	300
cacgtggagt	ttgctacagg	gagctgtca	cgatgaccca	gtggccccggg	gtctggatcc	360
cgatggcctc	ctgctcctcg	acgtgggtgt	caatggcg	gtccccggac	gagcctggct	420
gacgcagatc	tcagaatgtca	ggactttgaa	gaagcactac	gtgcaaaca	gggcctggcc	480
agctgttcgt	gggctccaca	cagcgcttct	tccagggcgg	cctccccctcg	ttcctacgt	540
gcaaccacag	catccagta	aacgcggccc	ggggccccc	gccccagctg	gtgcagcacc	600
tgccggcctc	agctatcagc	tcggcccttgc	atccagagc	cgaggccctg	cgcttccagc	660
tcgtacagc	cctgcaggcg	gaggagaacg	aggtcggtg	ccccggggc	tttagactgg	720
actcccgagg	acgttttgt	gtggatgtgg	acgagtgtgc	gtggatgtct	cacctctgcc	780
gagaggagaca	gcgctgtgt	aacctgtcg	ggtcttacc	ctgcctccccc	gactgtggc	840
ctggcttccg	ggtggctgtat	ggggccgct	gtaaaaatgt	ggacgaatgc	ctggaaagggg	900
ttggacgact	gtcaactacaa	ccagctctgc	gagaacaccc	caggcggtca	ccgctgcagc	960
tgccccaggg	gttaccggat	gcagggcccc	agcctgcct	gcttagatgt	caatgagtgc	1020
ctgcagctgc	ccaaggcctg	cgccattacc	tgccacaacc	tccagggcag	ctaccgctgc	1080
ctgtcccccc	caggccagac	cctcccttcg	gacggcaagg	ctgcaccc	actggagcgg	1140
aatggacaaa	atgtgaccac	cgtcagccac	cgaggccctc	tattgcctg	gtgcggccc	1200
tggcctcga	tccccgtac	ctccttaccac	gcctgggtct	ctcttcgtcc	gggtccccatg	1260
gcccctgag	gtgtggccg	ggctctgg	tctcctgggtt	tcatcaggca	gaacggagtc	1320
tgcaagacc	ttgacgagtg	ccgcgtgagg	aacctgtgtc	agcacgcctg	ccgcaacact	1380
gagggcagct	accagtgcct	gtgcccccc	ggctaccgtc	tgtctccccc	cgggaaagaa	1440
tgccaggaca	tcaacgagtg	cgaggaggag	agcatcgagt	gtggaccgg	ccagatgtgc	1500
ttcaacaccc	gtggcagcta	ccagtgtgt	gacacaccct	gtcctgccc	ctaccggcag	1560
ggcccccagcc	ctgggacgtg	cttccggc	tgctcgcagg	actgcggcac	gggcggccct	1620
tctacgctgc	agtaccggct	gctgccgt	ccccctggcg	tgcgcgccc	ccacgacgt	1680
gccccctca	ccgccttctc	cgaggtccgc	gtcccccca	accgcacccg	gctcagcatg	1740
ctggagcccg	accccttcgc	ctgcgtccgc	tgcgcgccc	ccccggccg	cttggccg	1800
gtctacaccc	gtgcgcgc	cacccttacc	ggctcaccgt	gctgctgcg	1860	
gcaccgcgccc	accaaaggct	cttcgtt	ctcatcgcc	tgtcccccta	ccccactaa	1920
acgggagagg	gcattggcg	ccgctctaga	ggatccct			1958

<210> 100

<211> 2444

<212> DNA

<213> Homo sapiens

<400> 100

ttacgc	caag ctggcacgag	caatgaaaga	gttaatctct	ttggctgggc	ctacagatga	60
catacagagt	acagtcccc	agg	ttcatgc	ttaaatatc	cttagagcat	120
tacgc	ctcg	ggagaaaata	ttatcc	tgttgcgtat	ggagctaa	180
gggttttaca	tcaccgg	tct	ggcagtg	aaattcatcc	acacttctct	240
gatcacaaga	attttggag	ttaaaaggc	aaaggatgaa	cattccaaa	caaata	300
gacaggaga	gat	tttct	ctcg	tttcc	tcaaacagt	360
ggaactgt	gcaata	actg	tagac	tatgg	ccaaatcg	420
gtttcttta	cttttgg	tgt	gagact	ctac	atccaa	480
tgctc	tgc	atgg	acttcc	cc	cgatggat	540
ccactccc	gt	gg	ctcg	ctt	gttgc	600
taataccatt	cg	aa	cttcc	catt	gttgc	660
aaaaccacat	tat	gg	acttcc	gttgc	actgacc	720
ctccaaaaca	cg	gg	acttcc	gttgc	actgac	780
ccaaagccaa	act	gg	gttgc	gttgc	actgttat	840
atattgat	tct	tt	acttgc	gttgc	gacaacc	900

cagttctgga gagtcttggc ttctggaaag aaattcaaag ggaatttatct caggatcaga	960
agctgataac gggattccct tgggccttca aggtgccagg cctgccccag tacctccaga	1020
gcctcaccag actagccatt gctgcagtgt gggccgcggc agccaagagt ggagagcggg	1080
agacaatgt ccccatctct ttctctcago tgtagatac tgccttcct gaagtgccgt	1140
cactaact ggaagccctc ttggaaaatg tcttagcage agactgttga ctggagaga	1200
aggcgctgcc acccttgcgt tgcaacatgg gagagaagtt ctatttgtt gccatgagg	1260
aaaatcaccc agaatgcctc tgcaagatatac tgaaaattct acactgcatt gaccctgg	1320
agtggctcc ccagacggag cactgtgtcc atctgacccc aaaggagttc ttgatctgga	1380
cgatggatat tgcttccat gaaagatctg aaattcagag ttagtctgt agacttgctt	1440
ccaaagtcat ttcccaccac atgcagacat gtgtggagaa cagggattt atagctgt	1500
agctgaagca gtgggttcag ctggtcatct tgcatagtga agaccatctt cctacagagt	1560
ctaggctggc cgtcggttgc gtcctcacca gtactacacc actttcctc accaacc	1620
atccattct tgagttgcag gatacacttgc ctctctggaa gtgtgtcctt acccttctgc	1680
agagtgagga gcaagctgtt agagatcgag ccacgaaac cgtgacaact gccatgtcac	1740
aagaaaatac ctggcagtc acagagttt ctttcgcac ggtggatgc tccatgc	1800
tggccctggc cctggccgc tcgtgtatc tgctccagca gtgggaccag ttggcccttg	1860
gactgcccatt ctgtctggga tggctgttgg gagagagtga tgacctcg	1920
agagcatgca tcaggtggaa gaagactacc tgtttggaaa agcagaagtc aacttttgg	1980
ccgagaccct gatctttgtt aaataccctt gcaaggaccc ttctgtctc ctctcaaagt	2040
ccggctggcg tcccccaagc cctgagatgc tctgtcaccc tcaaaggatg gtgtcagagc	2100
agtgcaccc ctgtctcag ttcttcagag agcttccacc agctgcttag tttgtgaaga	2160
cagtggagt cacaagacta cgcatcaag agggaaaggac ttggcttgc ttgaggctgc	2220
tggcctttt ggaaggaaag gaaggggaaag acacccatgt tctcagtggtt tggacttt	2280
atgcagaatc gaggcagttt actttccaa gaacagaagc ggcattgtga agaaaatctg	2340
ggggattttt atgggggtat gtgtggatcc ttcccttactt aatctgcag gaaacatgtt	2400
gaacataaat tcaaaaaattt tatccaaaaa aaaaaaaaaaaaaa aaaa	2444

<210> 101
<211> 2709
<212> DNA
<213> Homo sapiens

<400> 101	
ggcacgagat ttcttacagg tgaaacgcca tcattaggat tcactgttac gtttagtgcta	60
ttaaactcac tagcattttt attaatggcc gttatctaca ctaagctata ctgcacatgt	120
gaaaaagagg acctctcaga aaactcacaa tctagcatga ttaagcatgt cgcttggcta	180
atcttcacca attgcattt tttctgcctt gtggcggttt tttcatttgc accattgtatc	240
actgcacatct ctatcagccc cgaaataatg aagtctgtt ctctgatatt tttccatgc	300
ctgcttgcctt gaaatccagtc ctgtatgtt tttcaaccc aaagtttaaa gaagactgga	360
agttactgaa ggcacgtgtt accaagaaaa gtggatcagt ttctgttcc atcagtagcc	420
aaggctggatc tctggaaacagc gatttctat acgactgttg catgtactca catttgcagg	480
gcaacctgtac tgctgcac tgctgcgaat cgtttctttt aacaaagcca gtatcatgca	540
aacacttgat aaaatcacac agctgtctg cattggcagt ggcttcttgc caaagaccc	600
aggggctactg gtccgactgtt ggcacacatt cggcccaactc tgattatgca gatgaagaag	660
attcctttgtt ctcagacagt tctgaccagg tgcaggcctg tggacgagcc tgcttctacc	720
agagtagagg attccctttt gtgcgtatg cttacaatctt accaagatgtt aaagactgaa	780
ctactgtgtg tctaaccgtt tccccctgtca accaaaaatca gtgtttatag agtgaaccc	840
attctcatct ttcatctggg aagacttctt gtaatctactg cctgtgttca tttagaaga	900
ggaggggtgg cagtttattt ctcaaaccat tcattttcaaa agaaccagggtg cctaaattat	960
aaatgggtga aaaatgcaat gtccaaagca tgatgtatct gttgaaaca aatataatgac	1020
ttgaaaaggat tcttaggtgt agtagagacca tataatgtt gttttttgtt atccataaga	1080
agcaaaatttt taccttattt tttatgttgc acaagataaa gaacagctgt tataatttt	1140
taaaaaatttctt atttttttttt tttgtatgttca tataactgaa gaaaaatatac ttgtcaattt	1200
tacctaattt ttcatttcattt aatctcagga caacttactg cagggccaaa aaaggactg	1260
tcccaatgtt acctgtgaga gtatacatgtt gcattactttt attatgtttt cacttgc	1320
ccttgacata agagaactat aaatttttt taagcaattt ataaatctaa aacctgaa	1380
tgtttttttt acaatattaa cagctgttag gttaaaaaaa tagctggaca ttgtttttt	1440

65

gtcattatac	atgcgtttgg	tccaatcagt	aattttttct	taagtgtttt	gtgattacac	1500
tactagaaaa	aaaggtaaaag	gctaattgt	gtgtgggtt	agtcgatttg	gctaaactac	1560
taactaatgt	gggggtttaa	tagtatctga	gggatttgg	ggcttcatgt	aatgttctca	1620
ttaatgaata	cttccaataa	tcgttggctc	tactaatatt	ttccaatttg	ctgggatgtc	1680
accttagcaat	agcttggatt	atatagaaaag	taaactgtgg	tcaatacttg	catttaatta	1740
gacgaaacgg	ggagtaatta	tgacacgaaag	tacttaatgt	ttattttctta	gtgagctgga	1800
ttatcttgaa	cctgtgcata	taaatggaaa	tttccatataca	tcttccccat	actatTTTT	1860
ataaaagagc	ctattcaata	gctcagagggt	tgaactctgg	ttaaacaaga	taatatgtta	1920
ttaaaaaaaaaa	tagaagaaga	aaagataaaaag	cttagtccctg	tgtcttttaa	aaataaaaaa	1980
ttttacttgg	ttccccatct	atgggcattt	gaccattttac	tggttggagt	cttaaagtta	2040
taattgttca	atatgtttt	tgaacagtgt	gctaaatcaa	tagcaaacc	actgcccattat	2100
tagttattct	gaatatacta	aaaaaaatcca	gctagattgc	agtttaataa	ttaaactgtta	2160
catactgtgc	atataatgaa	tttttatctt	atgtaaattta	tttttagaac	acaagttggg	2220
aaatgtggct	tctgttcatt	tcgtttaatt	aaagctaccc	ccttaaactat	agtggctgcc	2280
agtagcagac	tgttaatttg	tgttttatat	actttttgca	ttgttaatag	tctttgttgt	2340
acattgtcag	tgtaaaaaaa	acagaatctt	tgtatataaa	aatcatgttag	tttgataaaa	2400
atgtggaaag	gattttatctt	cagtgtgttg	taatttgtt	aggccaaacta	tttacaaagg	2460
ttaaaaaattt	ctatcatgtt	tatccacaa	tctgataat	attaaatcat	aacttggtaa	2520
gaaactccct	attaaaaagg	ttttccaaa	attcagggtt	ttgaaaactt	ttcattttat	2580
tcattttaaa	actagaataa	cagatataata	aaagtgtttaa	tctttgtgt	atatggatgt	2640
aaatacaata	ttgtactcag	tgtttgaat	tatcaaagg	tctagaaagc	aaaaaaaaaa	2700
aaaaaaaaaa						2709

<210> 102
<211> 1722
<212> DNA
<213> *Homo sapiens*

<220>
<221> SITE
<222> (401)
<223> n equals a.t.q. or c

<220>
<221> SITE
<222> (695)
<223> n equals a,t,q, or c

ccgttcgcgca tgattttcta cccgcccggc	cgcgagggcg	cgctgctgct	ggcctcgta	1140
acgtggtcgg acgcggcgcc	acgttcgccc	ggcttgagcc	ggaaaggaggc	1200
gchgctcgacg acgtggcgcc	attgcacggg	cctgtcgctgc	gccagctctg	1260
ggcgtcgatca acgttgggc	ggaggaccag	cacagccagg	gtggctttgt	1320
ccggcgctct gcaaaccga	aaaggatag	tggacggtcc	cttatggccg	1380
gcccggcggc	acaccgccta	cccgcacggc	cggcggtcaa	1440
cgcggccca tcaagatcaa	cagccggaaag	gggcctgcat	cggacacggc	1500
gggcacgcata ctgacatgg	ggggcagggg	catgtgcata	gggtggccag	1560
catgacctgg caaaggaaga	aggcagccac	cctccagtcc	cagccccctcg	1620
aacacgaccc acacgaggac	ctcgattaa	agtatttcg	aaaaaaa	1680
aaaaaaaaaa aaaaaaaaaa	aaaaaaaaaa	aaaagggcgg	cc	1722

<210> 103

<211> 106

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (14)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (29)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 103

Met	Gly	Ser	Leu	Ser	Gly	Cys	Ala	Leu	Pro	Phe	Cys	Leu	Xaa	Val	Phe
1								5				10			15

Phe	Leu	Thr	Val	Ser	Pro	Ser	Ala	Val	Gly	Leu	Leu	Xaa	Phe	Ala	Gly
								20		25		30			

Gly	Pro	Leu	Gln	Thr	Leu	Phe	Ala	Trp	Val	Ser	Pro	Val	Glu	Ala	Ala
									35	40		45			

Glu	Gln	Gln	Arg	Leu	Leu	Pro	Val	Leu	Ser	Ser	Gly	Ser	Phe	Val	Ser
								50	55		60				

Glu	Gly	Thr	Cys	Gln	Met	Pro	Ala	Arg	Ala	Leu	Leu	Tyr	Glu	Val	Ser
								65	70		75		80		

Val	Gly	Pro	Tyr	Trp	Glu	Ile	Pro	Pro	Ser	Gln	Asp	Thr	Arg	Arg	Ser
								85		90		95			

Gly	Thr	Tyr	Leu	Arg	Arg	Gln	Ser	Asp	Pro
							100		105

<210> 104

<211> 86

<212> PRT

<213> Homo sapiens

<400> 104

Met	Thr	Leu	Pro	Ser	Arg	Ala	Leu	Ala	Ser	Leu	Gly	Val	Gly	Val	Trp
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

		67
1	5	10
		15
Gly Met Leu Arg Leu Asn Gln Val Thr Val Ser Cys Gly Gly Ser Arg		
20	25	30
Trp Ser Ser Arg Val Ala Leu Gly Ala Phe Ser Trp Val Cys Gly Val		
35	40	45
Ala Leu Val Leu Gln Pro Ser Gly Gly Gly Leu Gly Leu Thr Ser Pro		
.50	55	60
Ser Glu Gly Cys Trp Glu Gly Glu Leu Ala Leu Ala Val Leu Arg Ala		
65	70	75
80		
Pro Gly Gly Ser Pro Ser		
	85	
<210> 105		
<211> 302		
<212> PRT		
<213> Homo sapiens		
<400> 105		
Met Ala Arg Ala Arg Gly Ser Pro Cys Pro Pro Leu Pro Pro Gly Arg		
1	5	10
		15
Met Ser Trp Pro His Gly Ala Leu Leu Phe Leu Trp Leu Phe Ser Pro		
20	25	30
Pro Leu Gly Ala Gly Gly Gly Val Ala Val Thr Ser Ala Ala Gly		
35	40	45
Gly Gly Ser Pro Pro Ala Thr Ser Cys Pro Val Ala Cys Ser Cys Ser		
50	55	60
Asn Gln Ala Ser Arg Val Ile Cys Thr Arg Arg Asp Leu Ala Glu Val		
65	70	75
		80
Pro Ala Ser Ile Pro Val Asn Thr Arg Tyr Leu Asn Leu Gln Glu Asn		
85	90	95
Gly Ile Gln Val Ile Arg Thr Asp Thr Phe Lys His Leu Arg His Leu		
100	105	110
Glu Ile Leu Gln Leu Ser Lys Asn Leu Val Arg Lys Ile Glu Val Gly		
115	120	125
Ala Phe Asn Gly Leu Pro Ser Leu Asn Thr Leu Glu Leu Phe Asp Asn		
130	135	140
Arg Leu Thr Thr Val Pro Thr Gln Ala Phe Glu Tyr Leu Ser Lys Leu		
145	150	155
		160
Arg Glu Leu Trp Leu Arg Asn Asn Pro Ile Glu Ser Ile Pro Ser Tyr		
165	170	175
Ala Phe Asn Arg Val Pro Ser Leu Arg Arg Leu Asp Leu Gly Glu Leu		

68

180

185

190

Lys Arg Leu Glu Tyr Ile Ser Glu Ala Ala Phe Glu Gly Leu Val Asn
195 200 205

Leu Arg Tyr Leu Asn Leu Gly Met Cys Asn Leu Lys Asp Ile Pro Asn
210 215 220

Leu Thr Ala Leu Val Arg Leu Glu Glu Leu Glu Leu Ser Gly Asn Arg
225 230 235 240

Leu Asp Leu Ile Arg Pro Gly Ser Phe Gln Gly Leu Thr Ser Leu Arg
245 250 255

Lys Leu Trp Leu Met His Ala Gin Val Ala Thr Ile Glu Arg Asn Ala
260 265 270

Phe Asp Asp Leu Lys Ser Leu Glu Glu Leu Asn Leu Ser His Asn Asn
275 280 285

Leu Met Ser Leu Pro His Asp Leu Phe Thr Pro Leu His Arg
290 295 300

<210> 106
<211> 56
<212> PRT
<213> *Homo sapiens*

```
<220>
<221> SITE
<222> (56)
<223> Xaa equals stop translation
```

<400> 106
Met Pro Ser Ser Trp Leu Pro Gly Cys Phe Val Leu Leu Cys Leu Val
1 5 10 15

Ala Val Gly Cys Gln Leu Arg Glu Trp Gly Val Gly Gly Val Ser Ala
20 25 30

Val Gly Leu Leu Ala Leu Pro His Leu Gln Val Leu Gly Met Arg Gly
 35 40 45 .

Arg Gly Leu Ile Ser Gly Gly Xaa
50 55

<210> 107
<211> 189
<212> PRT
<213> *Homo sapiens*

<220>
<221> SITE
<222> (94)
<223> Xaa equals any of the naturally occurring L-amino acids

69

<400> 107
Met Ala Leu Leu Ser Arg Pro Ala Leu Thr Leu Leu Leu Leu Met
1 5 10 15

Ala Ala Val Val Arg Cys Gln Glu Gln Ala Gln Thr Thr Asp Trp Arg
20 25 30

Ala Thr Leu Lys Thr Ile Arg Asn Gly Val His Lys Ile Asp Thr Tyr
35 40 45

Leu Asn Ala Ala Leu Asp Leu Leu Gly Gly Glu Asp Gly Leu Cys Gln
50 55 60

Tyr Lys Cys Ser Asp Gly Ser Lys Pro Phe Pro Arg Tyr Gly Tyr Lys
65 70 75 80

Pro Ser Pro Pro Asn Gly Cys Gly Ser Pro Leu Phe Gly Xaa His Leu
85 90 95

Asn Ile Gly Ile Pro Ser Leu Thr Lys Cys Cys Asn Gln His Asp Arg
100 105 110

Cys Tyr Glu Thr Cys Gly Lys Ser Lys Asn Asp Cys Asp Glu Glu Phe
115 120 125

Gln Tyr Cys Leu Ser Lys Ile Cys Arg Asp Val Gln Lys Thr Leu Gly
130 135 140

Leu Thr Gln His Val Gln Ala Cys Glu Thr Thr Val Glu Leu Leu Phe
145 150 155 160

Asp Ser Val Ile His Leu Gly Cys Lys Pro Tyr Leu Asp Ser Gln Arg
165 170 175

Ala Ala Cys Arg Cys His Tyr Glu Glu Lys Thr Asp Leu
180 185

<210> 108
<211> 61
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (61)
<223> Xaa equals stop translation

<400> 108
Met Gly Asn Cys Gln Ala Gly His Asn Leu His Leu Cys Leu Ala His
1 5 10 15

His Pro Pro Leu Val Cys Ala Thr Leu Ile Leu Leu Leu Gly Leu
20 25 30

Ser Gly Leu Gly Leu Gly Ser Phe Leu Leu Thr His Arg Thr Gly Leu
35 40 45

70

Arg Thr Leu Thr Ser Pro Arg Thr Gly Ser Leu Phe Xaa
50 55 60

<210> 109
<211> 128
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (47)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (90)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 109
Met Arg Leu Glu Ser Leu Cys His Leu Cys Leu Ala Cys Leu Phe Phe
1 5 10 15

Arg Leu Pro Ala Thr Arg Thr Val Tyr Cys Met Asn Glu Ala Glu Ile
20 25 30

Val Asp Val Ala Leu Gly Ile Leu Ile Glu Ser Arg Lys Gln Xaa Lys
35 40 45

Ala Cys Glu Gln Pro Ala Leu Ala Gly Ala Asp Asn Pro Glu His Ser
50 55 60

Pro Pro Cys Ser Val Ser Pro His Thr Ser Ser Gly Ser Ser Ser Glu
65 70 75 80

Glu Glu Asp Ser Gly Lys Gln Ala Leu Xaa Pro Gly Leu Ser Pro Ser
85 90 95

Gln Arg Pro Gly Gly Ser Ser Ala Cys Ser Arg Ser Pro Glu Glu
100 105 110

Glu Glu Glu Glu Asp Val Leu Lys Tyr Val Arg Glu Ile Phe Phe Ser
115 120 125

<210> 110
<211> 69
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (50)
<223> Xaa equals any of the naturally occurring L-amino acids

71

<220>

<221> SITE

<222> (69)

<223> Xaa equals stop translation

<400> 110

Met Pro His Phe Leu Asp Trp Phe Val Pro Val Tyr Leu Val Ile Ser
1 5 10 15Val Leu Ile Leu Val Gly Phe Gly Ala Cys Ile Tyr Tyr Phe Glu Pro
20 25 30Gly Leu Gln Glu Ala His Lys Trp Arg Met Gln Arg Pro Leu Val Asp
35 40 45Arg Xaa Leu Arg Lys Thr Leu Met Val Arg Asp Asn Leu Ala Phe Gly
50 55 60Gly Pro Glu Val Xaa
65

<210> 111

<211> 123

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (123)

<223> Xaa equals stop translation

<400> 111

Met Ile Gly Gly Ile Thr Cys Ile Leu Ser Leu Ile Cys Ala Leu Ala
1 5 10 15Leu Ala Tyr Leu Asp Gln Arg Ala Glu Arg Ile Leu His Lys Glu Gln
20 25 30Gly Lys Thr Gly Glu Val Ile Lys Leu Thr Asp Val Lys Asp Phe Ser
35 40 45Leu Pro Leu Trp Leu Ile Phe Ile Ile Cys Val Cys Tyr Tyr Val Ala
50 55 60Val Phe Pro Phe Ile Gly Leu Gly Lys Val Phe Phe Thr Glu Lys Phe
65 70 75 80Gly Phe Ser Ser Gln Ala Ala Ser Ala Ile Asn Ser Val Val Tyr Val
85 90 95Ile Ser Ala Pro Met Ser Pro Val Phe Gly Leu Leu Val Asp Lys Thr
100 105 110Gly Lys Asn Ile Ile Trp Val Leu Cys Ala Xaa
115 120

72

<210> 112
<211> 83
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (83)
<223> Xaa equals stop translation

<400> 112
Met Glu Lys Gln Cys Cys Ser His Pro Val Ile Cys Ser Leu Ser Thr
1 5 10 15

Met Tyr Thr Phe Leu Leu Gly Ala Ile Phe Ile Ala Leu Ser Ser Ser
20 25 30

Arg Ile Leu Leu Val Lys Tyr Ser Ala Asn Glu Gly Lys Leu Arg Leu
35 40 45

Gly Ile Cys Met Glu His Phe His Leu Ile Thr His Leu Ser Leu Ala
50 55 60

Phe Gly Ser Val Ile Tyr Asn Met Glu Ile Ile Met Pro Phe Ala Ser
65 70 75 80

Cys Glu Xaa

<210> 113
<211> 345
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (53)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (345)
<223> Xaa equals stop translation

<400> 113
Met Asp Phe Leu Val Leu Phe Leu Phe Tyr Leu Ala Ser Val Leu Met
1 5 10 15

Gly Leu Val Leu Ile Cys Val Cys Ser Lys Thr His Ser Leu Lys Gly
20 25 30

Leu Ala Arg Gly Gly Ala Gln Ile Phe Ser Cys Ile Ile Pro Glu Cys
35 40 45

Leu Gln Arg Ala Xaa His Gly Leu Leu His Tyr Leu Phe His Thr Arg
50 55 60

73

Asn His Thr Phe Ile Val Leu His Leu Val Leu Gln Gly Met Val Tyr
65 70 75 80

Thr Glu Tyr Thr Trp Glu Val Phe Gly Tyr Cys Gln Glu Leu Glu Leu
85 90 95

Ser Leu His Tyr Leu Leu Leu Pro Tyr Leu Leu Leu Gly Val Asn Leu
100 105 110

Phe Phe Phe Thr Leu Thr Cys Gly Thr Asn Pro Gly Ile Ile Thr Lys
115 120 125

Ala Asn Glu Leu Leu Phe Leu His Val Tyr Glu Phe Asp Glu Val Met
130 135 140

Phe Pro Lys Asn Val Arg Cys Ser Thr Cys Asp Leu Arg Lys Pro Ala
145 150 155 160

Arg Ser Lys His Cys Ser Val Cys Asn Trp Cys Val His Arg Phe Asp
165 170 175

His His Cys Val Trp Val Asn Asn Cys Ile Gly Ala Trp Asn Ile Arg
180 185 190

Tyr Phe Leu Ile Tyr Val Leu Thr Leu Thr Ala Ser Ala Ala Thr Val
195 200 205

Ala Ile Val Ser Thr Thr Phe Leu Val His Leu Val Val Met Ser Asp
210 215 220

Leu Tyr Gln Glu Thr Tyr Ile Asp Asp Leu Gly His Leu His Val Met
225 230 235 240

Asp Thr Val Phe Leu Ile Gln Tyr Leu Phe Leu Thr Phe Pro Arg Ile
245 250 255

Val Phe Met Leu Gly Phe Val Val Val Leu Ser Phe Leu Leu Gly Gly
260 265 270

Tyr Leu Leu Phe Val Leu Tyr Leu Ala Ala Thr Asn Gln Thr Thr Asn
275 280 285

Glu Trp Tyr Arg Gly Asp Trp Ala Trp Cys Gln Arg Cys Pro Leu Val
290 295 300

Ala Trp Pro Pro Ser Ala Glu Pro Gln Val His Arg Asn Ile His Ser
305 310 315 320

His Gly Leu Arg Ser Asn Leu Gln Glu Ile Phe Leu Pro Ala Phe Pro
325 330 335

Cys His Glu Arg Lys Lys Gln Glu Xaa
340 345

<210> 114
<211> 181
<212> PRT

74

<213> Homo sapiens

<220>

<221> SITE

<222> (110)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 114

Met Ala Asp Pro His Val Ser Phe Leu Ser Phe Arg Gln Leu Phe Ser			
1	5	10	15

Trp Ala Ala Val Ile Leu Leu Arg Gly Ile Leu Gly Thr Val Ala Pro			
20	25	30	

Pro Pro Cys Pro Cys Val Leu Asp Leu Ala Val Tyr Pro Leu His Leu			
35	40	45	

Pro Val Glu Ala Pro Cys Leu Glu Val Val Phe Lys Gln Lys Asn Gly			
50	55	60	

Lys Asp Asn Cys Leu Val Phe Tyr Pro Asp Pro Ile Pro Leu Arg Gly			
65	70	75	80

Ser Leu Leu Gly Pro Phe Ile Lys Asn Gln Cys His Ser Ser Val Ile			
85	90	95	

Pro Leu Ser Asp Ser Ala Thr Ser Lys Ala Arg Ala Leu Xaa Leu Pro			
100	105	110	

Gly Arg Glu Thr Val Leu Ser Val Leu Pro Val Phe Ser Ser Pro Thr			
115	120	125	

Leu Pro Arg Thr His Ala Leu Gly Asp Ser Leu Gly Val Pro Gly Leu			
130	135	140	

Leu Val Cys Ser Glu Thr Ser Thr Leu Asn Asp His Trp Cys Cys Arg			
145	150	155	160

Arg Ala Gly Ala Tyr Ile Pro Ile Asn Arg Arg Phe Ser His Leu Met			
165	170	175	

Pro Leu Ala Phe Ser			
180			

<210> 115

<211> 116

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (116)

<223> Xaa equals stop translation

<400> 115

Met Pro Ser Ser Ser Ser Gly Leu Gly Ser Pro Ser Arg Pro Pro Ser			
1	5	10	15

75

Ser Phe Leu Cys Leu Leu Leu Leu Pro Pro Ala Ala Leu Ala
20 25 30

Leu Leu Leu Phe Phe Leu Asp Phe Phe Pro Pro Arg Ala Ala Val Ser
35 40 45

Pro Phe Leu Pro Asp His Cys Ser Ala Arg Gln Pro Arg Val Trp Arg
50 55 60

Arg Glu Thr Leu Asn Arg Ser Ala Ser Gly Leu Gly Cys Trp Ala Arg
65 70 75 80

Ser Thr Glu Gln Gly Ala Val Gly Val Ala Thr Gly Thr Val Leu Asp
85 90 95

Ile Ser Leu Pro Ala Ser Cys Leu Ser Leu Trp Pro Pro Gly Pro Ser
100 105 110

Gly Gly Ile Xaa
115

<210> 116
<211> 71
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (71)
<223> Xaa equals stop translation

<400> 116
Met Asn Leu Leu Gly Met Ile Phe Ser Met Cys Gly Leu Met Leu Lys
1 5 10 15

Leu Lys Trp Cys Ala Trp Val Ala Val Tyr Cys Ser Phe Ile Ser Phe
20 25 30

Ala Asn Ser Arg Ser Ser Glu Asp Thr Lys Gln Met Met Ser Ser Phe
35 40 45

Met Leu Ser Ile Ser Ala Val Val Met Ser Tyr Leu Gln Asn Pro Gln
50 55 60

Pro Met Thr Pro Pro Trp Xaa
65 70

<210> 117
<211> 64
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (64)

76

<223> Xaa equals stop translation

<400> 117

Met Arg Asp Leu Ser Phe Leu Tyr Thr Leu Leu Trp Leu Pro Glu Ile

1 5 10 15

Trp Gln Ala Leu Ala Gly Gly Ile Arg Leu Asp Glu Val Glu Leu Leu
 20 25 30

Glu Asn Glu Ala Val Leu Gly Glu Glu Met Arg Leu Tyr Arg Lys Ile
· 35 · 40 · 45

Asn Glu Val Val Leu Ser Gly Asn Glu Val Val Leu Gly Gly Lys Xaa
50 55 60

<210> 118

<211> 335

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (335)

<223> Xaa equals stop translation

<400> 118

Met Gly Ile Phe Pro Gly Ile Ile Leu Ile Phe Leu Arg Val Lys Phe
 1 5 10 15

Ala Thr Ala Ala Val Ile Val Ser Gly Val Ser Lys His Leu His Cys
20 25 30

Ile Ser His Gln Lys Ser Thr Thr Val Ser His Glu Met Ser Gly Leu
35 40 45

Asn	Trp	Lys	Pro	Phe	Val	Tyr	Gly	Gly	Leu	Ala	Ser	Ile	Val	Ala	Glu	
						50							55			60

Phe Gly Thr Phe Pro Val Asp Leu Thr Lys Thr Arg Leu Gln Val Gln
65 70 75 80

Gly Gln Ser Ile Asp Ala Arg Phe Lys Glu Ile Lys Tyr Arg Gly Met
85 90 95

Phe His Ala Leu Phe Arg Ile Cys Lys Glu Glu Gly Val Leu Ala Leu
 100 105 110

Tyr Ser Gly Ile Ala Pro Ala Leu Leu Arg Gln Ala Ser Tyr Gly Thr
115 120 125

Ile Lys Ile Gly Ile Tyr Gln Ser Leu Lys Arg Leu Phe Val Glu Arg
130 135 140

Leu Glu Asp Glu Thr Leu Leu Ile Asn Met Ile Cys Gly Val Val Ser

77

145	150	155	160
Gly Val Ile Ser Ser Thr Ile Ala Asn Pro Thr Asp Val Leu Lys Ile			
165	170	175	
Arg Met Gln Ala Gln Gly Ser Leu Phe Gln Gly Ser Met Ile Gly Ser			
180	185	190	
Phe Ile Asp Ile Tyr Gln Gln Glu Gly Thr Arg Gly Leu Trp Arg Gly			
195	200	205	
Val Val Pro Thr Ala Gln Arg Ala Ala Ile Val Val Gly Val Glu Leu			
210	215	220	
Pro Val Tyr Asp Ile Thr Lys Lys His Leu Ile Leu Ser Gly Met Met			
225	230	235	240
Gly Asp Thr Ile Leu Thr His Phe Val Ser Ser Phe Thr Cys Gly Leu			
245	250	255	
Ala Gly Ala Leu Ala Ser Asn Pro Val Asp Val Val Arg Thr Arg Met			
260	265	270	
Met Asn Gln Arg Ala Ile Val Gly His Val Asp Leu Tyr Lys Gly Thr			
275	280	285	
Val Asp Gly Ile Leu Lys Met Trp Lys His Glu Gly Phe Phe Ala Leu			
290	295	300	
Tyr Lys Gly Phe Trp Pro Asn Trp Leu Arg Leu Gly Pro Trp Asn Ile			
305	310	315	320
Ile Phe Phe Ile Thr Tyr Glu Gln Leu Lys Arg Leu Gln Ile Xaa			
325	330	335	
<210> 119			
<211> 221			
<212> PRT			
<213> Homo sapiens			
<220>			
<221> SITE			
<222> (51)			
<223> Xaa equals any of the naturally occurring L-amino acids			
<400> 119			
Met Ala Leu Ala Leu Ala Ala Leu Ala Ala Val Glu Pro Ala Cys Gly			
1	5	10	15
Ser Arg Tyr Gln Gln Leu Gln Asn Glu Glu Glu Ser Gly Glu Pro Glu			
20	25	30	
Gln Ala Ala Gly Asp Ala Pro Pro Pro Tyr Ser Ser Ile Ser Ala Glu			
35	40	45	
Ser Ala Xaa Tyr Phe Asp Tyr Lys Asp Glu Ser Gly Phe Pro Lys Pro			
50	55	60	

Pro Ser Tyr Asn Val Ala Thr Thr Leu Pro Ser Tyr Asp Glu Ala Glu
 65 70 75 80

 Arg Thr Lys Ala Glu Ala Thr Ile Pro Leu Val Pro Gly Arg Asp Glu
 85 90 95

 Asp Phe Val Gly Arg Asp Asp Phe Asp Asp Ala Asp Gln Leu Arg Ile
 100 105 110

 Gly Asn Asp Gly Ile Phe Met Leu Thr Phe Phe Met Ala Phe Leu Phe
 115 120 125

 Asn Trp Ile Gly Phe Phe Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala
 130 135 140

 Gly Arg Tyr Gly Ala Ile Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp
 145 150 155 160

 Ile Leu Ile Val Arg Phe Ser Thr Tyr Phe Pro Gly Tyr Phe Asp Gly
 165 170 175

 Gln Tyr Trp Leu Trp Trp Val Phe Leu Val Leu Gly Phe Leu Leu Phe
 180 185 190

 Leu Arg Gly Phe Ile Asn Tyr Ala Lys Val Arg Lys Met Pro Glu Thr
 195 200 205

 Phe Ser Asn Leu Pro Arg Thr Arg Val Leu Phe Ile Tyr
 210 215 220

<210> 120
<211> 473
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (473)
<223> Xaa equals stop translation

<400> 120
Met Lys Phe Leu Ile Phe Ala Phe Phe Gly Gly Val His Leu Leu Ser
1 5 10 15

Leu Cys Ser Gly Lys Ala Ile Cys Lys Asn Gly Ile Ser Lys Arg Thr
20 25 30

Phe Glu Glu Ile Lys Glu Glu Ile Ala Ser Cys Gly Asp Val Ala Lys
35 40 45

Ala Ile Ile Asn Leu Ala Val Tyr Gly Lys Ala Gln Asn Arg Ser Tyr
50 55 60

Glu Arg Leu Ala Leu Leu Val Asp Thr Val Gly Pro Arg Leu Ser Gly
65 70 75 80

79

Ser Lys Asn Leu Glu Lys Ala Ile Gln Ile Met Tyr Gln Asn Leu Gln
 85 90 95

Gln Asp Gly Leu Glu Lys Val His Leu Glu Pro Val Arg Ile Pro His
 100 105 110

Trp Glu Arg Gly Glu Glu Ser Ala Val Met Leu Glu Pro Arg Ile His
 115 120 125

Lys Ile Ala Ile Leu Gly Leu Gly Ser Ser Ile Gly Thr Pro Pro Glu
 130 135 140

Gly Ile Thr Ala Glu Val Leu Val Val Thr Ser Phe Asp Glu Leu Gln
 145 150 155 160

Arg Arg Ala Ser Glu Ala Arg Gly Lys Ile Val Val Tyr Asn Gln Pro
 165 170 175

Tyr Ile Asn Tyr Ser Arg Thr Val Gln Tyr Arg Thr Gln Gly Ala Val
 180 185 190

Glu Ala Ala Lys Val Gly Ala Leu Ala Ser Leu Ile Arg Ser Val Ala
 195 200 205

Ser Phe Ser Ile Tyr Ser Pro His Thr Gly Ile Gln Glu Tyr Gln Asp
 210 215 220

Gly Val Pro Lys Ile Pro Thr Ala Cys Ile Thr Val Glu Asp Ala Glu
 225 230 235 240

Met Met Ser Arg Met Ala Ser His Gly Ile Lys Ile Val Ile Gln Leu
 245 250 255

Lys Met Gly Ala Lys Thr Tyr Pro Asp Thr Asp Ser Phe Asn Thr Val
 260 265 270

Ala Glu Ile Thr Gly Ser Lys Tyr Pro Glu Gln Val Val Leu Val Ser
 275 280 285

Gly His Leu Asp Ser Trp Asp Val Gly Gln Gly Ala Met Asp Asp Gly
 290 295 300

Gly Gly Ala Phe Ile Ser Trp Glu Ala Leu Ser Leu Ile Lys Asp Leu
 305 310 315 320

Gly Leu Arg Pro Lys Arg Thr Leu Arg Leu Val Leu Trp Thr Ala Glu
 325 330 335

Glu Gln Gly Val Gly Ala Phe Gln Tyr Tyr Gln Leu His Lys Val
 340 345 350

Asn Ile Ser Asn Tyr Ser Leu Val Met Glu Ser Asp Ala Gly Thr Phe
 355 360 365

Leu Pro Thr Gly Leu Gln Phe Thr Gly Ser Glu Lys Ala Arg Ala Ile
 370 375 380

Met Glu Glu Val Met Ser Leu Leu Gln Pro Leu Asn Ile Thr Gln Val

385	390	395	400
Leu Ser His Gly Glu Gly Thr Asp Ile Asn Phe Trp Ile Gln Ala Gly			
	405	410	415
Val Pro Gly Ala Ser Leu Leu Asp Asp Leu Tyr Lys Tyr Phe Phe Phe			
	420	425	430
His His Ser His Gly Asp Thr Met Thr Val Met Asp Pro Lys Gln Met			
	435	440	445
Asn Val Ala Ala Ala Val Trp Ala Val Val Ser Tyr Val Val Ala Asp			
	450	455	460
Met Glu Glu Met Leu Pro Arg Ser Xaa			
	465	470	
<210> 121			
<211> 168			
<212> PRT			
<213> Homo sapiens			
<220>			
<221> SITE			
<222> (168)			
<223> Xaa equals stop translation			
<400> 121			
Met Ala Thr Leu Trp Gly Gly Leu Leu Arg Leu Gly Ser Leu Leu Ser			
1	5	10	15
Leu Ser Cys Leu Ala Leu Ser Val Leu Leu Leu Ala His Cys Gln Thr			
	20	25	30
Pro Pro Ser Asp Cys Leu His Val Val Glu Pro Met Pro Val Arg Gly			
35	40	45	
Pro Asp Val Glu Ala Tyr Cys Leu Arg Cys Glu Cys Lys Tyr Glu Glu			
50	55	60	
Arg Ser Ser Val Thr Ile Lys Val Thr Ile Ile Ile Tyr Leu Ser Ile			
65	70	75	80
Leu Gly Leu Leu Leu Leu Tyr Met Val Tyr Leu Thr Leu Val Glu Pro			
	85	90	95
Ile Leu Lys Arg Arg Leu Phe Gly His Ala Gln Leu Ile Gln Ser Asp			
	100	105	110
Asp Asp Ile Gly Asp His Gln Pro Phe Ala Asn Ala His Asp Val Leu			
	115	120	125
Ala Arg Ser Arg Ser Arg Ala Asn Val Leu Asn Lys Val Glu Tyr Ala			
	130	135	140
Gln Gln Arg Trp Lys Leu Gln Val Gln Glu Gln Arg Lys Ser Val Phe			
	145	150	155
			160

Asp Arg His Val Val Leu Ser Xaa
165

<210> 122
<211> 47
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (47)
<223> Xaa equals stop translation

<400> 122
Met Lys Phe Ile Leu Trp Arg Arg Phe Arg Trp Ala Ile Ile Leu Phe
1 5 10 15

Ile Ile Leu Phe Ile Leu Leu Phe Leu Ala Ile Phe Ile Tyr Ala
20 25 30

Phe Pro Asn Tyr Ala Ala Met Lys Leu Val Lys Pro Phe Ser Xaa
35 40 45

<210> 123
<211> 108
<212> PRT
<213> Homo sapiens

<400> 123
Met His Gln Asp Trp Leu Cys Asn Leu Gly Trp Pro Leu Leu Ser Leu
1 5 10 15

Trp Ala Ala Glu Ser Ala Pro His Val Ala Met Ala Ser Ala Thr Ala
20 25 30

Gln Leu Trp Ser Arg Pro Cys Gly Arg Thr His Met Val Ser Leu Ala
35 40 45

Leu Gly His Gln Glu Thr Gly Leu Trp Leu Cys Ser Ala Phe Gly Cys
50 55 60

Val Val Asp Ser Pro Trp Ala Ser Val Cys Pro Ser Val Lys Gly Gln
65 70 75 80

Leu Thr Val Cys Gly Ile Leu Pro Arg Val Pro Val Cys Val Tyr Val
85 90 95

Cys Ala Cys Val Arg Val Ser Met Cys Val His Ile
100 105

<210> 124
<211> 60
<212> PRT
<213> Homo sapiens

<400> 124
 Met Arg Gly Cys Val Pro Ala Phe Leu Leu His Val Leu Ser Leu Arg
 1 5 10 15
 Arg Ala Cys Cys Thr Gln Ala Ala Gln Val Phe Thr Ala Gln Leu Pro
 20 25 30
 Gly Arg Gln Val Ala Arg Arg Arg Gly Gly Trp His Glu Gln Gln Gly
 35 40 45
 Gly Pro Met Leu Cys Ser Ser His His Ser Arg Thr
 50 55 60

<210> 125
 <211> 248
 <212> PRT
 <213> Homo sapiens

<400> 125
 Met Ala Met Leu Pro Leu Val Leu His Trp Phe Phe Ile Glu Trp Tyr
 1 5 10 15
 Ser Gly Lys Lys Ser Ser Ala Leu Phe Gln His Ile Thr Ala Leu
 20 25 30
 Phe Glu Cys Ser Met Ala Ala Ile Ile Thr Leu Leu Val Ser Asp Pro
 35 40 45
 Val Gly Val Leu Tyr Ile Arg Ser Cys Arg Val Leu Met Leu Ser Asp
 50 55 60
 Trp Tyr Thr Met Leu Tyr Asn Pro Ser Pro Asp Tyr Val Thr Thr Val
 65 70 75 80
 His Cys Thr His Glu Ala Val Tyr Pro Leu Tyr Thr Ile Val Phe Ile
 85 90 95
 Tyr Tyr Ala Phe Cys Leu Val Met Met Leu Leu Arg Pro Leu Leu
 100 105 110
 Val Lys Lys Ile Ala Cys Gly Leu Gly Lys Ser Asp Arg Phe Lys Ser
 115 120 125
 Ile Tyr Ala Ala Leu Tyr Phe Phe Pro Ile Leu Thr Val Leu Gln Ala
 130 135 140
 Val Gly Gly Leu Leu Tyr Tyr Ala Phe Pro Tyr Ile Ile Leu Val
 145 150 155 160
 Leu Ser Leu Val Thr Leu Ala Val Tyr Met Ser Ala Ser Glu Ile Glu
 165 170 175
 Asn Cys Tyr Asp Leu Leu Val Arg Lys Lys Arg Leu Ile Val Leu Phe
 180 185 190
 Ser His Trp Leu Leu His Ala Tyr Gly Ile Ile Ser Ile Ser Arg Val

83

195 200 205

Asp Lys Leu Glu Gln Asp Leu Pro Pro Leu Ala Leu Val Pro Thr Pro
 210 215 220

Ala Leu Phe Tyr Leu Phe Thr Ala Lys Phe Thr Glu Pro Ser Arg Ile
 225 230 235 240

Leu Ser Glu Gly Ala Asn Gly His
 245

<210> 126

<211> 248

<212> PRT

<213> Homo sapiens

<400> 126

Met Glu Lys Ile Pro Glu Ile Gly Lys Phe Gly Glu Lys Ala Pro Pro
 1 5 10 15

Ala Pro Ser His Val Trp Arg Pro Ala Ala Leu Phe Leu Thr Leu Leu
 20 25 30

Cys Leu Leu Leu Leu Ile Gly Leu Gly Val Leu Ala Ser Met Phe His
 35 40 45

Val Thr Leu Lys Ile Glu Met Lys Lys Met Asn Lys Leu Gln Asn Ile
 50 55 60

Ser Glu Glu Leu Gln Arg Asn Ile Ser Leu Gln Leu Met Ser Asn Met
 65 70 75 80

Asn Ile Ser Asn Lys Ile Arg Asn Leu Ser Thr Thr Leu Gln Thr Ile
 85 90 95

Ala Thr Lys Leu Cys Arg Glu Leu Tyr Ser Lys Glu Gln Glu His Lys
 100 105 110

Cys Lys Pro Cys Pro Arg Arg Trp Ile Trp His Lys Asp Ser Cys Tyr
 115 120 125

Phe Leu Ser Asp Asp Val Gln Thr Trp Gln Glu Ser Lys Met Ala Cys
 130 135 140

Ala Ala Gln Asn Ala Ser Leu Leu Lys Ile Asn Asn Lys Asn Ala Leu
 145 150 155 160

Glu Phe Ile Lys Ser Gln Ser Arg Ser Tyr Asp Tyr Trp Leu Gly Leu
 165 170 175

Ser Pro Glu Glu Asp Ser Thr Arg Gly Met Arg Val Asp Asn Ile Ile
 180 185 190

Asn Ser Ser Ala Trp Val Ile Arg Asn Ala Pro Asp Leu Asn Asn Met
 195 200 205

Tyr Cys Gly Tyr Ile Asn Arg Leu Tyr Val Gln Tyr Tyr His Cys Thr

84

210

215

220

Tyr Lys Gln Arg Met Ile Cys Glu Lys Met Ala Asn Pro Val Gln Leu
 225 230 235 240
 Gly Ser Thr Tyr Phe Arg Glu Ala
 245

<210> 127
 <211> 612
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (245)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (246)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (249)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 127
 Met Ala Ala Ala Gly Arg Leu Pro Ser Ser Trp Ala Leu Phe Ser Pro
 1 5 10 15

Leu Leu Ala Gly Leu Ala Leu Leu Gly Val Gly Pro Val Pro Ala Arg
 20 25 30

Ala Leu His Asn Val Thr Ala Glu Leu Phe Gly Ala Glu Ala Trp Gly
 35 40 45

Thr Leu Ala Ala Phe Gly Asp Leu Asn Ser Asp Lys Gln Thr Asp Leu
 50 55 60

Phe Val Leu Arg Glu Arg Asn Asp Leu Ile Val Phe Leu Ala Asp Gln
 65 70 75 80

Asn Ala Pro Tyr Phe Lys Pro Lys Val Lys Val Ser Phe Lys Asn His
 85 90 95

Ser Ala Leu Ile Thr Ser Val Val Pro Gly Asp Tyr Asp Gly Asp Ser
 100 105 110

Gln Met Asp Val Leu Leu Thr Tyr Leu Pro Lys Asn Tyr Ala Lys Ser
 115 120 125

Glu Leu Gly Ala Val Ile Phe Trp Gly Gln Asn Gln Thr Leu Asp Pro
 130 135 140

Asn Asn Met Thr Ile Leu Asn Arg Thr Phe Gln Asp Glu Pro Leu Ile

85

145	150	155	160
Met Asp Phe Asn Gly Asp Leu Ile Pro Asp Ile Phe Gly Ile Thr Asn			
165		170	175
Glu Ser Asn Gln Pro Gln Ile Leu Leu Gly Gly Asn Leu Ser Trp His			
180	185		190
Pro Ala Leu Thr Thr Ser Lys Met Arg Ile Pro His Ser His Ala			
195		200	205
Phe Ile Asp Leu Thr Glu Asp Phe Thr Ala Asp Leu Phe Leu Thr Thr			
210	215		220
Leu Asn Ala Thr Thr Ser Thr Phe Gln Phe Glu Ile Trp Glu Asn Leu			
225	230	235	240
Asp Gly Asn Phe Xaa Xaa Ser Thr Xaa Leu Glu Lys Pro Gln Asn Met			
245		250	255
Met Val Val Gly Gln Ser Ala Phe Ala Asp Phe Asp Gly Asp Gly His			
260	265		270
Met Asp His Leu Leu Pro Gly Cys Glu Asp Lys Asn Cys Gln Lys Ser			
275		280	285
Thr Ile Tyr Leu Val Arg Ser Gly Met Lys Gln Trp Val Pro Val Leu			
290	295		300
Gln Asp Phe Ser Asn Lys Gly Thr Leu Trp Gly Phe Val Pro Phe Val			
305	310	315	320
Asp Glu Gln Gln Pro Thr Glu Ile Pro Ile Pro Ile Thr Leu His Ile			
325		330	335
Gly Asp Tyr Asn Met Asp Gly Tyr Pro Asp Ala Leu Val Ile Leu Lys			
340		345	350
Asn Thr Ser Gly Ser Asn Gln Gln Ala Phe Leu Leu Glu Asn Val Pro			
355	360		365
Cys Asn Asn Ala Ser Cys Glu Glu Ala Arg Arg Met Phe Lys Val Tyr			
370	375		380
Trp Glu Leu Thr Asp Leu Asn Gln Ile Lys Asp Ala Met Val Ala Thr			
385	390	395	400
Phe Phe Asp Ile Tyr Glu Asp Gly Ile Leu Asp Ile Val Val Leu Ser			
405		410	415
Lys Gly Tyr Thr Lys Asn Asp Phe Ala Ile His Thr Leu Lys Asn Asn			
420	425		430
Phe Glu Ala Asp Ala Tyr Phe Val Lys Val Ile Val Leu Ser Gly Leu			
435	440		445
Cys Ser Asn Asp Cys Pro Arg Lys Ile Thr Pro Phe Gly Val Asn Gln			
450	455	460	

Pro Gly Pro Tyr Ile Met Tyr Thr Thr Val Asp Ala Asn Gly Tyr Leu
465 470 475 480

Lys Asn Gly Ser Ala Gly Gln Leu Ser Gln Ser Ala His Leu Ala Leu
485 490 495

Gln Leu Pro Tyr Asn Val Leu Gly Leu Gly Arg Ser Ala Asn Phe Leu
500 505 510

Asp His Leu Tyr Val Gly Ile Pro Arg Pro Ser Gly Glu Lys Ser Ile
515 520 525

Arg Lys Gln Glu Trp Thr Ala Ile Ile Pro Asn Ser Gln Leu Ile Val
530 535 540

Ile Pro Tyr Pro His Asn Val Pro Arg Ser Trp Ser Ala Lys Leu Tyr
545 550 555 560

Leu Thr Pro Ser Asn Ile Val Leu Leu Thr Ala Ile Ala Leu Ile Gly
565 570 575

Val Cys Val Phe Ile Leu Ala Ile Ile Gly Ile Leu His Trp Gln Glu
580 585 590

Lys Lys Ala Asp Asp Arg Glu Lys Arg Gln Glu Ala His Arg Phe His
595 600 605

Phe Asp Ala Met
610

<210> 128

<211> 447

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (8)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (28)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (309)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (333)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 128

87

Met Glu Leu Ser Gln Met Ser Xaa Leu Met Gly Leu Ser Val Leu Leu
 1 5 10 15

Gly Leu Leu Ala Leu Met Ala Thr Ala Ala Val Xaa Arg Gly Trp Leu
 20 25 30

Arg Ala Gly Glu Glu Arg Ser Gly Arg Pro Ala Cys Gln Lys Ala Asn
 35 40 45

Gly Phe Pro Pro Asp Lys Ser Ser Gly Ser Lys Lys Gln Lys Gln Tyr
 50 55 60

Gln Arg Ile Arg Lys Glu Lys Pro Gln Gln His Asn Phe Thr His Arg
 65 70 75 80

Leu Leu Ala Ala Ala Leu Lys Ser His Ser Gly Asn Ile Ser Cys Met
 85 90 95

Asp Phe Ser Ser Asn Gly Lys Tyr Leu Ala Thr Cys Ala Asp Asp Arg
 100 105 110

Thr Ile Arg Ile Trp Ser Thr Lys Asp Phe Leu Gln Arg Glu His Arg
 115 120 125

Ser Met Arg Ala Asn Val Glu Leu Asp His Ala Thr Leu Val Arg Phe
 130 135 140

Ser Pro Asp Cys Arg Ala Phe Ile Val Trp Leu Ala Asn Gly Asp Thr
 145 150 155 160

Leu Arg Val Phe Lys Met Thr Lys Arg Glu Asp Gly Gly Tyr Thr Phe
 165 170 175

Thr Ala Thr Pro Glu Asp Phe Pro Lys Lys His Lys Ala Pro Val Ile
 180 185 190

Asp Ile Gly Ile Ala Asn Thr Gly Lys Phe Ile Met Thr Ala Ser Ser
 195 200 205

Asp Thr Thr Val Leu Ile Trp Ser Leu Lys Gly Gln Val Leu Ser Thr
 210 215 220

Ile Asn Thr Asn Gln Met Asn Asn Thr His Ala Ala Val Ser Pro Cys
 225 230 235 240

Gly Arg Phe Val Ala Ser Cys Gly Phe Thr Pro Asp Val Lys Val Trp
 245 250 255

Glu Val Cys Phe Gly Lys Lys Gly Glu Phe Gln Glu Val Val Arg Ala
 260 265 270

Phe Glu Leu Lys Gly His Ser Ala Ala Val His Ser Phe Ala Phe Ser
 275 280 285

Asn Asp Ser Arg Arg Met Ala Ser Val Ser Lys Asp Gly Thr Trp Lys
 290 295 300

Leu Trp Asp Thr Xaa Val Glu Tyr Lys Lys Gln Asp Pro Tyr Leu

88

305	310	315	320
Leu Lys Thr Gly Arg Phe Glu Glu Ala Ala Gly Ala Xaa Pro Cys Arg			
325		330	335
Leu Ala Leu Ser Pro Asn Ala Gln Val Leu Ala Leu Ala Ser Gly Ser			
340	345		350
Ser Ile His Leu Tyr Asn Thr Arg Arg Gly Glu Lys Glu Glu Cys Phe			
355		360	365
Glu Arg Val His Gly Glu Cys Ile Ala Asn Leu Ser Phe Asp Ile Thr			
370	375		380
Gly Arg Phe Leu Ala Ser Cys Gly Asp Arg Ala Val Arg Leu Phe His			
385	390	395	400
Asn Thr Pro Gly His Arg Ala Met Val Glu Glu Met Gln Gly His Leu			
405		410	415
Lys Arg Ala Ser Asn Glu Ser Thr Arg Gln Arg Leu Gln Gln Leu			
420	425		430
Thr Gln Ala Gln Glu Thr Leu Lys Ser Leu Gly Ala Leu Lys Lys			
435	440		445
<210> 129			
<211> 291			
<212> PRT			
<213> Homo sapiens			
<400> 129			
Met Leu Phe Leu Phe Ser Met Ala Thr Leu Leu Arg Thr Ser Phe Ser			
1	5	10	15
Asp Pro Gly Val Ile Pro Arg Ala Leu Pro Asp Glu Ala Ala Phe Ile			
20		25	30
Glu Met Glu Ile Glu Ala Thr Asn Gly Ala Val Pro Gln Gly Gln Arg			
35	40		45
Pro Pro Pro Arg Ile Lys Asn Phe Gln Ile Asn Asn Gln Ile Val Lys			
50	55	60	
Leu Lys Tyr Cys Tyr Thr Cys Lys Ile Phe Arg Pro Pro Arg Ala Ser			
65	70	75	80
His Cys Ser Ile Cys Asp Asn Cys Val Glu Arg Phe Asp His His Cys			
85		90	95
Pro Trp Val Gly Asn Cys Val Gly Lys Arg Asn Tyr Arg Tyr Phe Tyr			
100		105	110
Leu Phe Ile Leu Ser Leu Ser Leu Leu Thr Ile Tyr Val Phe Ala Phe			
115	120		125
Asn Ile Val Tyr Val Ala Leu Lys Ser Leu Lys Ile Gly Phe Leu Glu			

	⁸⁹	
130	135	140
Thr Leu Lys Glu Thr Pro Gly Thr Val Leu Glu Val Leu Ile Cys Phe		
145	150	155
Phe Thr Leu Trp Ser Val Val Gly Leu Thr Gly Phe His Thr Phe Leu		
165	170	175
Val Ala Leu Asn Gln Thr Thr Asn Glu Asp Ile Lys Gly Ser Trp Thr		
180	185	190
Gly Lys Asn Arg Val Gln Asn Pro Tyr Ser His Gly Asn Ile Val Lys		
195	200	205
Asn Cys Cys Glu Val Leu Cys Gly Pro Leu Pro Pro Ser Val Leu Asp		
210	215	220
Arg Arg Gly Ile Leu Pro Leu Glu Glu Ser Gly Ser Arg Pro Pro Ser		
225	230	240
Thr Gln Glu Thr Ser Ser Leu Leu Pro Gln Ser Pro Ala Pro Thr		
245	250	255
Glu His Leu Asn Ser Asn Glu Met Pro Glu Asp Ser Ser Thr Pro Glu		
260	265	270
Glu Met Pro Pro Pro Glu Pro Pro Glu Pro Pro Gln Glu Ala Ala Glu		
275	280	285
Ala Glu Lys		
290		
<210> 130		
<211> 78		
<212> PRT		
<213> Homo sapiens		
<400> 130		
Met Val Arg Lys Trp Leu Thr Phe Val Glu His Leu Leu Cys Ala Trp		
1	5	10
Pro Arg Leu Gly Ala Phe Val Pro Arg Val Thr Pro Ser Glu Cys Ser		
20	25	30
Ser Leu Pro His Ser Asn Trp Gly Val Gly Gly Arg Ala Ala Gln Leu		
35	40	45
Thr Gly Ala Glu Leu Lys Thr His Ser Trp Val Cys Leu Gly Trp Ala		
50	55	60
Val Leu Val Ala Pro Val Ala Asn Thr Arg Ala Pro Phe Thr		
65	70	75

<210> 131		
<211> 333		
<212> PRT		

90

<213> Homo sapiens

<220>

<221> SITE

<222> (97)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 131

Met	Leu	Met	Phe	Ala	Val	Ile	Val	Ala	Ser	Ser	Gly	Leu	Leu	Leu	Met
1															15

Ile	Glu	Arg	Gly	Ile	Leu	Ala	Glu	Met	Lys	Pro	Leu	Pro	Leu	His	Pro
															30
				20				25							

Pro	Gly	Arg	Glu	Gly	Thr	Ala	Trp	Arg	Gly	Lys	Ala	Pro	Lys	Pro	Gly
															45
				35				40							

Gly	Leu	Ser	Leu	Arg	Ala	Gly	Asp	Ala	Asp	Leu	Gln	Val	Arg	Gln	Asp
															60
				50				55							

Val	Arg	Asn	Arg	Thr	Leu	Arg	Ala	Val	Cys	Gly	Gln	Pro	Gly	Met	Pro
															80
				65				70			75				

Arg	Asp	Pro	Trp	Asp	Leu	Pro	Val	Gly	Gln	Arg	Arg	Thr	Leu	Leu	Arg
															95
				85				90							

Xaa	Ile	Leu	Val	Ser	Asp	Arg	Tyr	Arg	Phe	Leu	Tyr	Cys	Tyr	Val	Pro
															110
				100				105							

Lys	Val	Ala	Cys	Ser	Asn	Trp	Lys	Arg	Val	Met	Lys	Val	Leu	Ala	Gly
															125
				115				120							

Val	Leu	Asp	Ser	Val	Asp	Val	Arg	Leu	Lys	Met	Asp	His	Arg	Ser	Asp
															140
				130				135							

Leu	Val	Phe	Leu	Ala	Asp	Leu	Arg	Pro	Glu	Glu	Ile	Arg	Tyr	Arg	Leu
															160
				145				150			155				

Gln	His	Tyr	Phe	Lys	Phe	Leu	Phe	Val	Arg	Glu	Pro	Leu	Glu	Arg	Leu
															175
				165				170							

Leu	Ser	Ala	Tyr	Arg	Asn	Lys	Phe	Gly	Glu	Ile	Arg	Glu	Tyr	Gln	Gln
															190
				180				185							

Arg	Tyr	Gly	Ala	Glu	Ile	Val	Arg	Arg	Tyr	Arg	Ala	Gly	Ala	Gly	Pro
															205
				195				200							

Ser	Pro	Ala	Gly	Asp	Asp	Val	Thr	Phe	Pro	Glu	Phe	Leu	Arg	Tyr	Leu
															220
				210				215							

Val	Asp	Glu	Asp	Pro	Glu	Arg	Met	Asn	Glu	His	Trp	Met	Pro	Val	Tyr
															240
				225				230			235				

His	Leu	Cys	Gln	Pro	Cys	Ala	Val	His	Tyr	Asp	Phe	Val	Gly	Ser	Tyr
															255
				245				250							

Glu	Arg	Leu	Glu	Ala	Asp	Ala	Asn	Gln	Val	Leu	Glu	Trp	Val	Arg	Ala
															270
				260				265			270				

Pro Pro His Val Arg Phe Pro Ala Arg Gln Ala Trp Tyr Arg Pro Ala
 275 280 285
 Ser Pro Glu Ser Leu His Tyr His Leu Cys Ser Ala Pro Arg Ala Leu
 290 295 300
 Leu Gln Asp Val Leu Pro Lys Tyr Ile Leu Asp Phe Ser Leu Phe Ala
 305 310 315 320
 Tyr Pro Leu Pro Asn Val Thr Lys Glu Ala Cys Gln Gln
 325 330

 <210> 132
 <211> 164
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (126)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <400> 132
 Met Leu Pro Leu Leu Ile Ile Cys Leu Leu Pro Ala Ile Glu Gly Lys
 1 5 10 15

 Asn Cys Leu Arg Cys Trp Pro Glu Leu Ser Ala Leu Ile Asp Tyr Asp
 20 25 30

 Leu Gln Ile Leu Trp Val Thr Pro Gly Pro Pro Thr Glu Leu Ser Gln
 35 40 45

 Ser Ile His Ser Leu Phe Leu Glu Asp Asn Asn Phe Leu Lys Pro Trp
 50 55 60

 Tyr Leu Asp Arg Asp His Leu Glu Glu Glu Thr Ala Lys Phe Phe Thr
 65 70 75 80

 Gln Val His Gln Ala Ile Lys Thr Leu Arg Asp Asp Lys Thr Val Leu
 85 90 95

 Leu Glu Glu Ile Tyr Thr His Lys Asn Leu Phe Thr Glu Arg Leu Asn
 100 105 110

 Lys Ile Ser Asp Gly Leu Lys Glu Lys Gly Ala Pro Pro Xaa Ser Met
 115 120 125

 Asn Ala Phe Pro Ala Pro Ser Pro Thr Cys Thr Pro Glu Pro Leu Gly
 130 135 140

 Ser Val Cys Leu Pro Ser Thr Ser Val Ser Leu Pro Ser His Leu Pro
 145 150 155 160

 Gly Ser Leu Gln

<210> 133
<211> 245
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (245)
<223> Xaa equals stop translation

<400> 133
Met Val Ala Val Val Gly Val Tyr Ala Arg Leu Met Lys His Ala Glu Ala
1 5 10 15
Ala Leu Ala Cys Leu Ala Val Asp Pro Ala Ile Leu Leu Ile Val Val
20 25 30
Gly Val Leu Met Phe Leu Leu Thr Phe Cys Gly Cys Ile Gly Ser Leu
35 40 45
Arg Glu Asn Ile Cys Leu Leu Gln Thr Phe Ser Leu Cys Leu Thr Ala
50 55 60
Val Phe Leu Leu Gln Leu Ala Ala Gly Ile Leu Gly Phe Val Phe Ser
65 70 75 80
Asp Lys Ala Arg Gly Lys Val Ser Glu Ile Ile Asn Asn Ala Ile Val
85 90 95
His Tyr Arg Asp Asp Leu Asp Leu Gln Asn Leu Ile Asp Phe Gly Gln
100 105 110
Lys Lys Phe Ser Cys Cys Gly Gly Ile Ser Tyr Lys Asp Trp Ser Gln
115 120 125
Asn Met Tyr Phe Asn Cys Ser Glu Asp Asn Pro Ser Arg Glu Arg Cys
130 135 140
Ser Val Pro Tyr Ser Cys Cys Leu Pro Thr Pro Asp Gln Ala Val Ile
145 150 155 160
Asn Thr Met Cys Gly Gln Gly Met Gln Ala Phe Asp Tyr Leu Glu Ala
165 170 175
Ser Lys Val Ile Tyr Thr Asn Gly Cys Ile Asp Lys Leu Val Asn Trp
180 185 190
Ile His Ser Asn Leu Phe Leu Leu Gly Gly Val Ala Leu Gly Leu Ala
195 200 205
Ile Pro Gln Leu Val Gly Ile Leu Leu Ser Gln Ile Leu Val Asn Gln
210 215 220
Ile Lys Asp Gln Ile Lys Leu Gln Leu Tyr Asn Gln Gln His Arg Ala
225 230 235 240
Asp Pro Trp Tyr Xaa

93

245

<210> 134
<211> 56
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (56)
<223> Xaa equals stop translation

<400> 134
Met Gly Thr Val Gly Leu Trp Pro Ser Trp Leu Trp Leu Pro Ala Ser
1 5 10 15

Trp Pro Leu Thr Ser Cys Gly Val Thr Arg Arg Arg Leu Arg Gly Pro
20 25 30

Gly Leu Arg Arg Thr Ser Gln Thr Gly Arg His Thr Ser Pro Cys Pro
35 40 45

Thr Ala Thr Trp Ala Glu Ser Xaa
50 55

<210> 135
<211> 55
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (47)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (51)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (55)
<223> Xaa equals stop translation

<400> 135
Met Ser Ile Val Met Ser Pro Leu Leu Leu Pro Ile Cys Tyr Leu Asn
1 5 10 15

Leu Leu Leu Phe Phe Val Asn Leu Ala Lys Asn Leu Ser Ile Leu Phe
20 25 30

Val Ser Ser Lys Lys Tyr Thr Phe Val Phe Met Ile Ser Leu Xaa Phe
35 40 45

Phe His Xaa Tyr Phe Ile Xaa

94

50

55

<210> 136
 <211> 89
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (89)
 <223> Xaa equals stop translation

<400> 136
 Met Ala Ile Ile Ser Phe Glu Leu Leu Phe Leu Met Asn Leu Pro Thr
 1 5 10 15

Val Asn Ser Ser Asn Phe Lys Leu Ile Ile Pro Glu Asp Val Thr Leu
 20 25 30

Sér Phe Val Ser His Leu Asp Ile Thr Val Asn His Phe Val Phe Leu
 35 40 45

Ser Thr Phe Glu Leu Ala Gly Val Ile Glu Gly Lys Pro Leu Pro Asp
 50 55 60

Ser Lys Ser Asp Leu Cys Pro Ile Leu Gly Gln Leu Trp Phe His Ile
 65 70 75 80

Leu Leu Phe Phe Ile Phe Trp Val Xaa
 85

<210> 137
 <211> 62
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (62)
 <223> Xaa equals stop translation

<400> 137
 Met Arg Leu Pro Ile Ala Pro His Leu Gln Tyr Phe Met Trp Ser Val
 1 5 10 15

Leu Leu Phe Leu Val Ile Leu Val Asp Met Lys Trp His Leu Ser Val
 20 25 30

Ala Phe His Tyr Ile Ser Leu Met Thr Asn Gly Ile Leu Ser Pro Phe
 35 40 45

Gln Cys Leu Leu Ala Ile His Val Ser Leu Phe Phe Val Xaa
 50 55 60

<210> 138

95

<211> 106
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (106)
<223> Xaa equals stop translation

<400> 138
Met Cys Leu Leu Pro Gly Gly Val Leu Leu Ile Trp Ser Cys Ala Ser
1 5 10 15

Gly Thr Pro Ala Ser His Thr Lys Asp Trp Gly Arg Cys Lys Phe Ser
20 25 30

Ala Ala Thr Lys Arg Thr Ala Glu Ser Asn Leu Glu Ser Thr Gln Leu
35 40 45

Met Leu Ala Ser Gln Ile Asp Pro Leu Leu Ala Glu Cys Trp His Leu
50 55 60

Cys Ala Ser Val Ser Ser Val Asn Gly Gly Asp Lys Lys Cys Val
65 70 75 80

His Thr Ser Arg Ala Val Gly Arg Ile Lys Leu Cys Ser Asp Thr Ile
85 90 95

Arg Ala Cys Ser Gly Trp Tyr Leu Gln Xaa
100 105

<210> 139
<211> 52
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (52)
<223> Xaa equals stop translation

<400> 139
Met Ser His Ser Val Phe Ala His Tyr Ile Phe Asn Ile Leu Leu
1 5 10 15

Leu Leu Leu Leu Leu Ile Gly Phe Leu Tyr Ser Met Pro Phe Ile
20 25 30

Tyr Lys Asp Thr Lys Thr His Val Cys Asn Phe Asn Asn Ile Phe
35 40 45

Pro Ile Leu Xaa
50

<210> 140
<211> 119

96

<212> PRT
 <213> Homo sapiens

<400> 140
 Met Lys Trp Arg Arg Lys Ser Ala Tyr Trp Lys Ala Leu Lys Val Phe
 1 5 10 15
 Lys Leu Pro Val Glu Phe Leu Leu Leu Thr Val Pro Val Val Asp
 20 25 30
 Pro Asp Lys Asp Asp Gln Asn Trp Lys Arg Pro Leu Asn Cys Leu His
 35 40 45
 Leu Val Ile Ser Pro Leu Val Val Leu Thr Leu Gln Ser Gly Thr
 50 55 60
 Tyr Gly Val Tyr Glu Ile Gly Gly Leu Val Pro Val Trp Val Val Val
 65 70 75 80
 Val Ile Ala Gly Thr Ala Leu Ala Ser Val Thr Phe Phe Ala Thr Ser
 85 90 95
 Asp Ser Gln Pro Pro Arg Leu His Trp Leu Phe Ala Phe Leu Gly Phe
 100 105 110
 Leu Thr Ser Ala Leu Trp Ile
 115

<210> 141
 <211> 59
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (59)
 <223> Xaa equals stop translation

<400> 141
 Met Cys Ser Gly Ser Phe Lys Glu Leu Tyr Leu Val Pro Ile Ser Leu
 1 5 10 15
 Phe Ser Thr Cys Val Leu Gly Phe Tyr Phe His Asn Phe Leu Leu Leu
 20 25 30
 Ile Ile Leu Phe Ser Ile Leu Leu Arg Lys Ile Thr Gly Lys Leu Phe
 35 40 45
 Phe Thr Tyr Tyr His Phe Ser Cys Gly Val Xaa
 50 55

<210> 142
 <211> 100
 <212> PRT
 <213> Homo sapiens

97

<220>

<221> SITE

<222> (100)

<223> Xaa equals stop translation

<400> 142

Met	Leu	Phe	Phe	Leu	Ser	Leu	Phe	Leu	Ser	Leu	Leu	Leu	Thr	Leu	Ser
1															15

Leu	Pro	Ser	Phe	Leu	Pro	Phe	Ser	Phe	Phe	Phe	Ser	Leu	Phe	Pro
														30
20								25						

His	Leu	Ser	Ala	Cys	Leu	Leu	Pro	Ser	Leu	Pro	Ser	Pro	Pro	Phe	Pro
35								40				45			

Leu	Pro	Pro	Ser	Leu	Pro	Ser	Phe	Leu	Pro	Ser	Phe	Leu	Pro	Ser	Phe
50							55				60				

Leu	Pro	Ser	Leu	Leu	Ser	Pro	Ser	Phe	Pro	Ala	Phe	Phe	Pro	Ser	Phe
65							70			75		80			

Cys	Gln	Leu	Ala	Arg	Arg	Ser	Pro	Arg	Lys	Ser	Thr	Gln	Met	Leu	Gln
							85		90			95			

Ser	Thr	Ser	Xaa
			100

<210> 143

<211> 65

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (61)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (65)

<223> Xaa equals stop translation

<400> 143

Met	Ala	Val	Leu	Leu	Ile	Thr	Ile	Leu	Leu	Phe	Leu	Cys	Leu	Gly	Tyr
1															15

Tyr	Arg	Val	Ile	Thr	Glu	Ile	Ser	Arg	Lys	Thr	Pro	Ala	Cys	Arg	Met
20								25				30			

Phe	Thr	Ser	Ser	Leu	Ser	Ser	Trp	Tyr	Ile	Met	Arg	Lys	Leu	Tyr	Asp
35								40				45			

Thr	Pro	Gly	Glu	Val	Phe	Leu	Ser	His	Ala	Ile	Val	Xaa	Phe	Leu	Lys
50								55			60				

Xaa

65

<210> 144
<211> 67
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (67)
<223> Xaa equals stop translation

<400> 144
Met Leu Asn Gln Pro Cys Ile Leu Gly Met Lys Pro Thr Trp Leu Trp
1 5 10 15

Trp Ile Ser Phe Leu Met Cys Cys Trp Val Trp Leu Ala Ser Val Leu
20 25 30

Leu Gly Ile Phe Ala Ser Ile Phe Ile Arg Asp Ile Gly Leu Glu Phe
35 40 45

Ser Phe Phe Val Met Cys Leu Pro Gly Phe Gly Ile Arg Val Met Leu
50 55 60

Ala Ser Xaa
65

<210> 145
<211> 59
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (59)
<223> Xaa equals stop translation

<400> 145
Met Thr Ala Met Ser Ile His Leu Phe Cys Thr Ala Leu Ser Cys Gly
1 5 10 15

Ser Ser Gly Gln Cys Asn Lys Ala Ile Lys Arg Asn Lys Ile Ser Asn
20 25 30

Asp Trp Lys Asp Val Asn Val Ser Ser Phe Ile Glu Asn Met Ile His
35 40 45

Arg Tyr Thr Tyr Thr Asn Ala Leu Asn Ser Xaa
50 55

<210> 146
<211> 56
<212> PRT
<213> Homo sapiens

99

<220>
<221> SITE
<222> (56)
<223> Xaa equals stop translation

<400> 146
Met Ser His Cys Thr Trp Pro Val Cys Leu Phe Cys Leu Val Pro Pro
1 5 10 15

Pro Met Gly Asp Leu Lys Glu Val Cys Leu Pro His Arg Cys Pro Gly
20 25 30

Arg Thr Ala Cys Cys Ser Tyr Ser Glu Pro His Leu Gln Thr Glu Glu
35 40 45

Asp Arg Arg Thr Leu Ile Cys Xaa
50 55

<210> 147
<211> 66
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (66)
<223> Xaa equals stop translation

<400> 147
Met Thr Asn Gly His Gln Val Leu Leu Leu Leu Leu Thr Ser Ala
1 5 10 15

Val Ala Ala Gly Pro Trp Pro Gln Val His Ala Gly Gln Trp Gly Trp
20 25 30

Met Cys Leu Pro Pro Gly Leu Pro Ser Val Gln Ala Arg Ser Gly Leu
35 40 45

Gly Gly Leu Pro Gly Gly Pro Gln Trp Val Pro Gly Gly Ala Arg Gly
50 55 60

Tyr Xaa
65

<210> 148
<211> 328
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (328)
<223> Xaa equals stop translation

<400> 148
Met Ala Cys Arg Lys Leu Ala Val Ala His Pro Leu Leu Leu Arg

		100			
1	5	10			
His Leu Pro Met Ile Ala Ala Leu Leu His Gly Arg Thr His Leu Asn					
		20	25	30	
Phe Gln Glu Phe Arg Gln Gln Asn His Leu Ser Cys Phe Leu His Val					
		35	40	45	
Leu Gly Leu Leu Glu Leu Leu Gln Pro His Val Phe Arg Ser Glu His					
		50	55	60	
Gln Gly Ala Leu Trp Asp Cys Leu Leu Ser Phe Ile Arg Leu Leu Leu					
		65	70	75	80
Asn Tyr Arg Lys Ser Ser Arg His Leu Ala Ala Phe Ile Asn Lys Phe					
		85	90	95	
Val Gln Phe Ile His Lys Tyr Ile Thr Tyr Asn Ala Pro Ala Ala Ile					
		100	105	110	
Ser Phe Leu Gln Lys His Ala Asp Pro Leu His Asp Leu Ser Phe Asp					
		115	120	125	
Asn Ser Asp Leu Val Met Leu Lys Ser Leu Leu Ala Gly Leu Ser Leu					
		130	135	140	
Pro Ser Arg Asp Asp Arg Thr Asp Arg Gly Leu Asp Glu Glu Gly Glu					
		145	150	155	160
Glu Glu Ser Ser Ala Gly Ser Leu Pro Leu Val Ser Val Ser Leu Phe					
		165	170	175	
Thr Pro Leu Thr Ala Ala Glu Met Ala Pro Tyr Met Lys Arg Leu Ser					
		180	185	190	
Arg Gly Gln Thr Val Glu Asp Leu Leu Glu Val Leu Ser Asp Ile Asp					
		195	200	205	
Glu Met Ser Arg Arg Pro Glu Ile Leu Ser Phe Phe Ser Thr Asn					
		210	215	220	
Leu Gln Arg Leu Met Ser Ser Ala Glu Glu Cys Cys Arg Asn Leu Ala					
		225	230	235	240
Phe Ser Leu Ala Leu Arg Ser Met Gln Asn Ser Pro Ser Ile Ala Ala					
		245	250	255	
Ala Phe Leu Pro Thr Phe Met Tyr Cys Leu Gly Ser Gln Asp Phe Glu					
		260	265	270	
Val Val Gln Thr Ala Leu Arg Asn Leu Pro Glu Tyr Ala Leu Leu Cys					
		275	280	285	
Gln Glu His Ala Ala Val Leu Leu His Arg Ala Phe Leu Val Gly Met					
		290	295	300	
Tyr Gly Gln Met Asp Pro Ser Ala Gln Ile Ser Glu Ala Leu Arg Ile					
		305	310	315	320

101

Leu His Met Glu Ala Val Met Xaa
325

<210> 149
<211> 90
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (10)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (13)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (90)
<223> Xaa equals stop translation

<400> 149
Met Gly Phe Leu Gln Leu Leu Val Val Xaa Val Leu Xaa Ser Glu His
1 5 10 15

Arg Val Ala Gly Ala Ala Glu Val Phe Gly Asn Ser Ser Glu Gly Leu
20 25 30

Ile Glu Phe Ser Val Gly Lys Phe Arg Tyr Phe Glu Leu Asn Arg Pro
35 40 45

Phe Pro Glu Glu Ala Ile Leu His Asp Ile Ser Ser Asn Val Thr Phe
50 55 60

Leu Ile Phe Gln Ile His Ser Gln Tyr Gln Asn Thr Thr Val Ser Phe
65 70 75 80

Ser Pro Arg Arg Ser Pro Thr Met Xaa
85 90

<210> 150
<211> 149
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (149)
<223> Xaa equals stop translation

<400> 150
Met Ala Gly Ser Pro Leu Leu Trp Gly Pro Arg Ala Gly Gly Val Gly
1 5 10 15

102

Leu Leu Val Leu Leu Leu Gly Leu Phe Arg Pro Pro Pro Ala Leu
20 25 30

Cys Ala Arg Pro Val Lys Glu Pro Arg Gly Leu Ser Ala Ala Ser Pro
35 40 45

Pro Leu Ala Arg Leu Ala Leu Ala Ala Ser Gly Gly Gln Cys Pro
50 55 60

Glu Val Arg Arg Arg Gly Arg Cys Arg Pro Gly Ala Gly Ala Gly Ala
65 70 75 80

Ser Ala Gly Ala Glu Arg Gln Glu Arg Ala Arg Ala Glu Ala Gln Arg
85 90 95

Leu Arg Ile Ser Arg Arg Ala Ser Trp Arg Ser Cys Cys Ala Ser Gly
100 105 110

Ala Pro Pro Ala Thr Leu Ile Arg Leu Trp Ala Trp Thr Thr Thr Pro
115 120 125

Thr Arg Leu Gln Arg Ser Ser Leu Ala Leu Cys Ser Ala Pro Ala Leu
130 135 140

Thr Leu Pro Pro Xaa
145

<210> 151
<211> 391
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (391)
<223> Xaa equals stop translation

<400> 151
Met Leu Pro Thr Trp Trp Ile Val Ser Ser Trp Leu Val Trp Gly Val
1 5 10 15

Ile Leu Phe Val Tyr Leu Val Ile Arg Ala Leu Arg Leu Trp Arg Thr
20 25 30

Ala Lys Leu Gln Val Thr Leu Lys Lys Tyr Ser Val His Leu Glu Asp
35 40 45

Met Ala Thr Asn Ser Arg Ala Phe Thr Asn Leu Val Arg Lys Ala Leu
50 55 60

Arg Leu Ile Gln Glu Thr Glu Val Ile Ser Arg Gly Phe Thr Leu Val
65 70 75 80

Ser Ala Ala Cys Pro Phe Asn Lys Ala Gly Gln His Pro Ser Gln His
85 90 95

103

Leu Ile Gly Leu Arg Lys Ala Val Tyr Arg Thr Leu Arg Ala Asn Phe
100 105 110

Gln Ala Ala Arg Leu Ala Thr Leu Tyr Met Leu Lys Asn Tyr Pro Leu
115 120 125

Asn Ser Glu Ser Asp Asn Val Thr Asn Tyr Ile Cys Val Val Pro Phe
130 135 140

Lys Glu Leu Gly Leu Gly Leu Ser Glu Glu Gln Ile Ser Glu Glu Glu
145 150 155 160

Ala His Asn Phe Thr Asp Gly Phe Ser Leu Pro Ala Leu Lys Val Leu
165 170 175

Phe Gln Leu Trp Val Ala Gln Ser Ser Glu Phe Phe Arg Arg Leu Ala
180 185 190

Leu Leu Leu Ser Thr Ala Asn Ser Pro Pro Gly Pro Leu Leu Thr Pro
195 200 205

Ala Leu Leu Pro His Arg Ile Leu Ser Asp Val Thr Gln Gly Leu Pro
210 215 220

His Ala His Ser Ala Cys Leu Glu Glu Leu Lys Arg Ser Tyr Glu Phe
225 230 235 240

Tyr Arg Tyr Phe Glu Thr Gln His Gln Ser Val Pro Gln Cys Leu Ser
245 250 255

Lys Thr Gln Gln Lys Ser Arg Glu Leu Asn Asn Val His Thr Ala Val
260 265 270

Arg Ser Leu Gln Leu His Leu Lys Ala Leu Leu Asn Glu Val Ile Ile
275 280 285

Leu Glu Asp Glu Leu Glu Lys Leu Val Cys Thr Lys Glu Thr Gln Glu
290 295 300

Leu Val Ser Glu Ala Tyr Pro Ile Leu Glu Gln Lys Leu Lys Leu Ile
305 310 315 320

Gln Pro His Val Gln Ala Ser Asn Asn Cys Trp Glu Glu Ala Ile Ser
325 330 335

Gln Val Asp Lys Leu Leu Arg Arg Asn Thr Asp Lys Lys Gly Lys Pro
340 345 350

Glu Ile Ala Cys Glu Asn Pro His Cys Thr Val Ser Thr Phe Glu Ala
355 360 365

Ala Tyr Ser Thr His Cys Arg Gln Arg Ser Asn Pro Arg Gly Ala Gly
370 375 380

Ile Arg Ser Leu Cys Arg Xaa
385 390

104

<210> 152
 <211> 99
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (99)
 <223> Xaa equals stop translation

<400> 152
 Met Thr Thr Arg Gln Pro Thr Ala Val Ser Trp Pro Cys Trp Leu Met
 1 5 10 15

Ser Ser Ser Leu Ser Thr Ala Cys Leu Ala Trp Thr Leu Thr Gly Ser
 20 25 30

Leu Ala Arg Glu Ala Thr Arg Arg Ala Arg Ser Leu Ser Pro Thr Trp
 35 40 45

Asn Cys Ser Ala Arg Gln Val Pro Pro Ser Pro Pro His Ser Gly Leu
 50 55 60

Gly Arg Arg Gly Trp Ala His Cys His Leu Thr Cys Leu Leu Val Thr
 65 70 75 80

Gln Leu Phe Arg Val Gly Arg Ile His Pro Ile Leu Ser Leu Pro Leu
 85 90 95

Val Thr Xaa

<210> 153
 <211> 61
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (61)
 <223> Xaa equals stop translation

<400> 153
 Met Ser His Cys Ala Arg Pro Thr Phe Leu Thr Leu Leu Ala Ser
 1 5 10 15

Cys Phe Trp Ala Ala Ala Ile Pro Asn Arg Asn Val Ile Leu Ser Val
 20 25 30

Ser Phe Arg Pro Leu His Met Gln Phe Thr Leu Ser Ile Leu Val Phe
 35 40 45

Ile Leu Arg Ile Leu Ile Leu Arg Ser Phe Leu Xaa
 50 55 60

<210> 154

105

<211> 393

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (393)

<223> Xaa equals stop translation

<400> 154

Met	Glu	Trp	Trp	Ala	Ser	Ser	Pro	Leu	Arg	Leu	Trp	Leu	Leu	Phe
1								5					10	

Leu	Leu	Pro	Ser	Ala	Gln	Gly	Arg	Gln	Lys	Glu	Ser	Gly	Ser	Lys	Trp
								20		25			30		

Lys	Val	Phe	Ile	Asp	Gln	Ile	Asn	Arg	Ser	Leu	Glu	Asn	Tyr	Glu	Pro
								35		40			45		

Cys	Ser	Ser	Gln	Asn	Cys	Ser	Cys	Tyr	His	Gly	Val	Ile	Glu	Glu	Asp
								50		55		60			

Leu	Thr	Pro	Phe	Arg	Gly	Gly	Ile	Ser	Arg	Lys	Met	Met	Ala	Glu	Val
65							70			75			80		

Val	Arg	Arg	Lys	Leu	Gly	Thr	His	Tyr	Gln	Ile	Thr	Lys	Asn	Arg	Leu
						85			90			95			

Tyr	Arg	Glu	Asn	Asp	Cys	Met	Phe	Pro	Ser	Arg	Cys	Ser	Gly	Val	Glu
						100		105				110			

His	Phe	Ile	Leu	Glu	Val	Ile	Gly	Arg	Leu	Pro	Asp	Met	Glu	Met	Val
						115		120			125				

Ile	Asn	Val	Arg	Asp	Tyr	Pro	Gln	Val	Pro	Lys	Trp	Met	Glu	Pro	Ala
						130		135			140				

Ile	Pro	Val	Phe	Ser	Phe	Ser	Lys	Thr	Ser	Glu	Tyr	His	Asp	Ile	Met
145						150			155			160			

Tyr	Pro	Ala	Trp	Thr	Phe	Trp	Glu	Gly	Gly	Pro	Ala	Val	Trp	Pro	Ile
						165		170			175				

Tyr	Pro	Thr	Gly	Leu	Gly	Arg	Trp	Asp	Leu	Phe	Arg	Glu	Asp	Leu	Val
				180		185			185			190			

Arg	Ser	Ala	Ala	Gln	Trp	Pro	Trp	Lys	Lys	Lys	Asn	Ser	Thr	Ala	Tyr
						195		200			205				

Phe	Arg	Gly	Ser	Arg	Thr	Ser	Pro	Glu	Arg	Asp	Pro	Leu	Ile	Leu	Leu
						210		215			220				

Ser	Arg	Lys	Asn	Pro	Lys	Leu	Val	Asp	Ala	Glu	Tyr	Thr	Lys	Asn	Gln
225						225		230		235			240		

Ala	Trp	Lys	Ser	Met	Lys	Asp	Thr	Leu	Gly	Lys	Pro	Ala	Ala	Lys	Asp
						245		250			255				

106

Val His Leu Val Asp His Cys Lys Tyr Lys Tyr Leu Phe Asn Phe Arg
 260 265 270

Gly Val Ala Ala Ser Phe Arg Phe Lys His Leu Phe Leu Cys Gly Ser
 275 280 285

Leu Val Phe His Val Gly Asp Glu Trp Leu Glu Phe Phe Tyr Pro Gln
 290 295 300

Leu Lys Pro Trp Val His Tyr Ile Pro Val Lys Thr Asp Leu Ser Asn
 305 310 315 320

Val Gln Glu Leu Leu Gln Phe Val Lys Ala Asn Asp Asp Val Ala Gln
 325 330 335

Glu Ile Ala Glu Arg Gly Ser Gln Phe Ile Arg Asn His Leu Gln Met
 340 345 350

Asp Asp Ile Thr Cys Tyr Trp Glu Asn Leu Leu Ser Glu Tyr Ser Lys
 355 360 365

Phe Leu Ser Tyr Asn Val Thr Arg Arg Lys Gly Tyr Asp Gln Ile Ile
 370 375 380

Pro Lys Met Leu Lys Thr Glu Leu Xaa
 385 390

<210> 155

<211> 75

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (75)

<223> Xaa equals stop translation

<400> 155

Met Leu Ile Leu Phe Leu Ser Val Cys Leu Phe Val Phe Leu Leu Thr
 1 5 10 15

Val Arg Ala Leu Cys Cys Arg Ser Ala Gly Val Trp Leu Arg Ser Thr
 20 25 30

Pro Asp Pro Val Cys Leu Gly Phe Ala Arg Gly Gly Cys Arg Ile Ala
 35 40 45

Met Ile Ala Ala Cys Phe Ser Ser Gly Ser Phe Val Pro Glu Gly His
 50 55 60

Pro Pro Asp Ala Ser Gln Ser Ser Pro Val Xaa
 65 70 75

<210> 156

<211> 82

<212> PRT

107

<213> Homo sapiens

<220>

<221> SITE

<222> (82)

<223> Xaa equals stop translation

<400> 156

Met	Trp	Pro	Leu	Leu	Ala	Val	Ser	Pro	Phe	Gly	Leu	Val	Trp	Ala	Ser
1														15	

Ser	Gln	Ser	Gly	Ser	Leu	Leu	Leu	Arg	Ala	Ser	Ile	Pro	Arg	Gln	His
					20			25					30		

Ser	Arg	Arg	Ala	Trp	His	Phe	Tyr	Ser	Glu	Val	Trp	Gln	Ser	His	Ser
					35			40				45			

Val	Ala	Ser	Val	Leu	Leu	Tyr	Leu	Leu	Val	Arg	Ala	Ile	Thr	Lys	Met
						50			55			60			

Cys	Ile	Gly	Ser	Lys	Lys	Arg	Asp	Ile	Thr	Pro	Thr	Thr	Arg	Trp	Lys
					65		70			75			80		

Lys Xaa

<210> 157

<211> 54

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (49)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (54)

<223> Xaa equals stop translation

<400> 157

Met	Ser	His	His	Ala	Gly	Leu	Gly	Gly	Ile	Leu	Phe	Ser	Leu	Lys
1						5			10				15	

Ile	Ser	Phe	Phe	Ile	Ala	Leu	Ala	Val	Val	Gly	Gly	Ser	Arg	Gly	Val
						20			25			30			

Asn	Asp	Cys	Gln	Leu	Gly	Gly	Cys	Arg	Val	Gly	Ser	Cys	Pro	Arg	Val
						35		40			45				

Xaa Val Arg Val Ala Xaa

50

<210> 158

<211> 103

108

<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (103)
<223> Xaa equals stop translation

<400> 158

Met Thr Val Arg Arg Leu Ser Leu Leu Cys Arg Asp Leu Trp Ala Leu
1 5 10 15

Trp Leu Leu Leu Lys Ala Gly Ala Val Arg Gly Ala Arg Ala Gly Pro
20 25 30

Arg Leu Pro Gly Arg Cys Cys Gly Ala Thr Cys Gly Asp Ala Gly Arg
35 40 45

Gly Trp Thr Phe Trp Ala Gln Pro Cys Pro Gln Lys Leu Leu Gly Gln
50 55 60

Lys Pro Gly Ala Gly Cys Arg Gly Trp Val Leu Gly Trp Val Pro
65 70 75 80

Pro Arg Pro Glu Glu Pro Cys Ser Leu Ala Gly Lys Val Cys Thr Gly
85 90 95

Leu Ala Arg Trp Met Val Xaa
100.

<210> 159
<211> 575
<212> PRT
<213> Homo sapiens

<400> 159

Met Arg Val Leu Val Val Thr Ile Ala Pro Ile Tyr Trp Ala Leu Ala
1 5 10 15

Arg Glu Ser Gly Glu Ala Leu Asn Gly His Ser Leu Thr Gly Gly Lys
20 25 30

Phe Arg Gln Glu Ser His Val Glu Phe Ala Thr Gly Glu Leu Leu Thr
35 40 45

Met Thr Gln Trp Pro Gly Val Trp Ile Pro Met Ala Ser Cys Ser Ser
50 55 60

Thr Trp Trp Ser Met Ala Leu Ser Pro Asp Ser Leu Ala Asp Ala Asp
65 70 75 80

Leu Gln Val Gln Asp Phe Glu Glu His Tyr Val Gln Thr Gly Pro Gly
85 90 95

Gln Leu Phe Val Gly Ser Thr Gln Arg Phe Phe Gln Gly Gly Leu Pro
100 105 110

109

Ser Phe Leu Arg Cys Asn His Ser Ile Gln Tyr Asn Ala Ala Arg Gly
115 120 125

Pro Gln Pro Gln Leu Val Gln His Leu Arg Ala Ser Ala Ile Ser Ser
130 135 140

Ala Phe Asp Pro Glu Ala Glu Ala Leu Arg Phe Gln Leu Ala Thr Ala
145 150 155 160

Leu Gln Ala Glu Glu Asn Glu Val Gly Cys Pro Glu Gly Phe Glu Leu
165 170 175

Asp Ser Gln Gly Ala Phe Cys Val Asp Val Asp Glu Cys Ala Trp Asp
180 185 190

Ala His Leu Cys Arg Glu Gly Gln Arg Cys Val Asn Leu Leu Gly Ser
195 200 205

Tyr Arg Cys Leu Pro Asp Cys Gly Pro Gly Phe Arg Val Ala Asp Gly
210 215 220

Ala Gly Cys Glu Asp Val Asp Glu Cys Leu Glu Gly Leu Asp Asp Cys
225 230 235 240

His Tyr Asn Gln Leu Cys Glu Asn Thr Pro Gly Gly His Arg Cys Ser
245 250 255

Cys Pro Arg Gly Tyr Arg Met Gln Gly Pro Ser Leu Pro Cys Leu Asp
260 265 270

Val Asn Glu Cys Leu Gln Leu Pro Lys Ala Cys Ala Tyr Gln Cys His
275 280 285

Asn Leu Gln Gly Ser Tyr Arg Cys Leu Cys Pro Pro Gly Gln Thr Leu
290 295 300

Leu Arg Asp Gly Lys Ala Cys Thr Ser Leu Glu Arg Asn Gly Gln Asn
305 310 315 320

Val Thr Thr Val Ser His Arg Gly Pro Leu Leu Pro Trp Leu Arg Pro
325 330 335

Trp Ala Ser Ile Pro Gly Thr Ser Tyr His Ala Trp Val Ser Leu Arg
340 345 350

Pro Gly Pro Met Ala Leu Ser Ser Val Gly Arg Ala Trp Cys Pro Pro
355 360 365

Gly Phe Ile Arg Gln Asn Gly Val Cys Thr Asp Leu Asp Glu Cys Arg
370 375 380

Val Arg Asn Leu Cys Gln His Ala Cys Arg Asn Thr Glu Gly Ser Tyr
385 390 395 400

Gln Cys Leu Cys Pro Ala Gly Tyr Arg Leu Leu Pro Ser Gly Lys Asn
405 410 415

Cys Gln Asp Ile Asn Glu Cys Glu Glu Ser Ile Glu Cys Gly Pro

		110
420	425	430
Gly Gln Met Cys Phe Asn Thr Arg Gly Ser Tyr Gln Cys Val Asp Thr		
435	440	445
Pro Cys Pro Ala Thr Tyr Arg Gln Gly Pro Ser Pro Gly Thr Cys Phe		
450	455	460
Arg Arg Cys Ser Gln Asp Cys Gly Thr Gly Gly Pro Ser Thr Leu Gln		
465	470	475
Tyr Arg Leu Leu Pro Leu Pro Leu Gly Val Arg Ala His His Asp Val		
485	490	495
Ala Arg Leu Thr Ala Phe Ser Glu Val Gly Val Pro Ala Asn Arg Thr		
500	505	510
Glu Leu Ser Met Leu Glu Pro Asp Pro Arg Ser Pro Phe Ala Leu Arg		
515	520	525
Pro Leu Arg Ala Gly Leu Gly Ala Val Tyr Thr Arg Arg Ala Leu Thr		
530	535	540
Arg Ala Gly Leu Tyr Arg Leu Thr Val Arg Ala Ala Ala Pro Arg His		
545	550	555
Gln Ser Val Phe Val Leu Leu Ile Ala Val Ser Pro Tyr Pro Tyr		
565	570	575
<210> 160		
<211> 643		
<212> PRT		
<213> Homo sapiens		
<400> 160		
Met Gly Glu Pro Asn Arg His Pro Ser Met Phe Leu Leu Leu Val		
1	5	10
Leu Glu Arg Leu Tyr Ala Ser Pro Met Asp Gly Thr Ser Ser Ala Leu		
20	25	30
Ser Met Gly Pro Phe Val Pro Phe Ile Met Arg Cys Gly His Ser Pro		
35	40	45
Val Tyr His Ser Arg Glu Met Ala Ala Arg Ala Leu Val Pro Phe Val		
50	55	60
Met Ile Asp His Ile Pro Asn Thr Ile Arg Thr Leu Leu Ser Thr Leu		
65	70	75
Pro Ser Cys Thr Asp Gln Cys Phe Arg Gln Asn His Ile His Gly Thr		
85	90	95
Leu Leu Gln Val Phe His Leu Leu Gln Ala Tyr Ser Asp Ser Lys His		
100	105	110
Gly Thr Asn Ser Asp Phe Gln His Glu Leu Thr Asp Ile Thr Val Cys		

111

115	120	125
Thr Lys Ala Lys Leu Trp Leu Ala Lys Arg Gln Asn Pro Cys Leu Val 130	135	140
Thr Arg Ala Val Tyr Ile Asp Ile Leu Phe Leu Leu Thr Cys Cys Leu 145	150	155
Asn Arg Ser Ala Lys Asp Asn Gln Pro Val Leu Glu Ser Leu Gly Phe 165	170	175
Trp Glu Glu Val Arg Gly Ile Ile Ser Gly Ser Glu Leu Ile Thr Gly 180	185	190
Phe Pro Trp Ala Phe Lys Val Pro Gly Leu Pro Gln Tyr Leu Gln Ser 195	200	205
Leu Thr Arg Leu Ala Ile Ala Ala Val Trp Ala Ala Ala Lys Ser 210	215	220
Gly Glu Arg Glu Thr Asn Val Pro Ile Ser Phe Ser Gln Leu Leu Glu 225	230	235
Ser Ala Phe Pro Glu Val Arg Ser Leu Thr Leu Glu Ala Leu Leu Glu 245	250	255
Lys Phe Leu Ala Ala Ser Gly Leu Gly Glu Lys Gly Val Pro Pro 260	265	270
Leu Leu Cys Asn Met Gly Glu Lys Phe Leu Leu Leu Ala Met Lys Glu 275	280	285
Asn His Pro Glu Cys Phe Cys Lys Ile Leu Lys Ile Leu His Cys Met 290	295	300
Asp Pro Gly Glu Trp Leu Pro Gln Thr Glu His Cys Val His Leu Thr 305	310	315
Pro Lys Glu Phe Leu Ile Trp Thr Met Asp Ile Ala Ser Asn Glu Arg 325	330	335
Ser Glu Ile Gln Ser Val Ala Leu Arg Leu Ala Ser Lys Val Ile Ser 340	345	350
His His Met Gln Thr Cys Val Glu Asn Arg Glu Leu Ile Ala Ala Glu 355	360	365
Leu Lys Gln Trp Val Gln Leu Val Ile Leu Ser Cys Glu Asp His Leu 370	375	380
Pro Thr Glu Ser Arg Leu Ala Val Val Glu Val Leu Thr Ser Thr Thr 385	390	395
Pro Leu Phe Leu Thr Asn Pro His Pro Ile Leu Glu Leu Gln Asp Thr 405	410	415
Leu Ala Leu Trp Lys Cys Val Leu Thr Leu Leu Gln Ser Glu Glu Gln 420	425	430

112

Ala Val Arg Asp Ala Ala Thr Glu Thr Val Thr Ala Met Ser Gln
 435 440 445
 Glu Asn Thr Cys Gln Ser Thr Glu Phe Ala Phe Cys Gln Val Asp Ala
 450 455 460
 Ser Ile Ala Leu Ala Leu Ala Val Leu Cys Asp Leu Leu Gln
 465 470 475 480
 Gln Trp Asp Gln Leu Ala Pro Gly Leu Pro Ile Leu Leu Gly Trp Leu
 485 490 495
 Leu Gly Glu Ser Asp Asp Leu Val Ala Cys Val Glu Ser Met His Gln
 500 505 510
 Val Glu Glu Asp Tyr Leu Phe Glu Lys Ala Glu Val Asn Phe Trp Ala
 515 520 525
 Glu Thr Leu Ile Phe Val Lys Tyr Leu Cys Lys His Leu Phe Cys Leu
 530 535 540
 Leu Ser Lys Ser Gly Trp Arg Pro Pro Ser Pro Glu Met Leu Cys His
 545 550 555 560
 Leu Gln Arg Met Val Ser Glu Gln Cys His Leu Leu Ser Gln Phe Phe
 565 570 575
 Arg Glu Leu Pro Pro Ala Ala Glu Phe Val Lys Thr Val Glu Phe Thr
 580 585 590
 Arg Leu Arg Ile Gln Glu Glu Arg Thr Leu Ala Cys Leu Arg Leu Leu
 595 600 605
 Ala Phe Leu Glu Gly Lys Glu Gly Glu Asp Thr Leu Val Leu Ser Val
 610 615 620
 Trp Asp Ser Tyr Ala Glu Ser Arg Gln Leu Thr Leu Pro Arg Thr Glu
 625 630 635 640
 Ala Ala Cys

<210> 161
 <211> 191
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (191)
 <223> Xaa equals stop translation

<400> 161
 Met Ser Ser Gly Thr Glu Leu Leu Trp Pro Gly Ala Ala Leu Leu Val
 1 5 10 15

113

Leu Leu Gly Val Ala Ala Ser Leu Cys Val Arg Cys Ser Arg Pro Gly
 20 25 30

Ala Lys Arg Ser Glu Lys Ile Tyr Gln Gln Arg Ser Leu Arg Glu Asp
 35 40 45

Gln Gln Ser Phe Thr Gly Ser Arg Thr Tyr Ser Leu Val Gly Gln Ala
 50 55 60

Trp Pro Gly Pro Leu Ala Asp Met Ala Pro Thr Arg Lys Asp Lys Leu
 65 70 75 80

Leu Gln Phe Tyr Pro Ser Leu Glu Asp Pro Ala Ser Ser Arg Tyr Gln
 85 90 95

Asn Phe Ser Lys Gly Ser Arg His Gly Ser Glu Glu Ala Tyr Ile Asp
 100 105 110

Pro Ile Ala Met Glu Tyr Tyr Asn Trp Gly Arg Phe Ser Lys Pro Pro
 115 120 125

Glu Asp Asp Asp Ala Asn Ser Tyr Glu Asn Val Leu Ile Cys Lys Gln
 130 135 140

Lys Thr Thr Glu Thr Gly Ala Gln Gln Glu Gly Ile Gly Gly Leu Cys
 145 150 155 160

Arg Gly Asp Leu Ser Leu Ser Leu Ala Leu Lys Thr Gly Pro Thr Ser
 165 170 175

Gly Leu Cys Pro Ser Ala Ser Pro Glu Glu Asp Glu Gly Ile Xaa
 180 185 190

<210> 162
 <211> 64
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (64)
 <223> Xaa equals stop translation

<400> 162
 Met Lys His Val Leu Asn Leu Tyr Leu Leu Gly Val Val Leu Thr Leu
 1 5 10 15

Leu Ser Ile Phe Val Arg Val Met Glu Ser Leu Glu Gly Leu Leu Glu
 20 25 30

Ser Pro Ser Pro Gly Thr Ser Trp Thr Thr Arg Ser Gln Leu Ala Asn
 35 40 45

Thr Glu Pro Thr Lys Gly Leu Pro Asp His Pro Ser Arg Ser Met Xaa
 50 55 60

<210> 163
<211> 118
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (118)
<223> Xaa equals stop translation

<400> 163
Met Ile Phe Leu Thr Val Leu Pro Leu Ala Phe Leu Phe Leu His Ser
1 5 10 15

Gly Phe Tyr His Tyr Ile Ser Phe Ser Cys Leu Phe Ser Leu Ser Leu
20 25 30

Ala Leu Phe Phe Leu Asp Val Ala Thr Phe Arg Arg Pro Gly Gln
35 40 45

Leu Phe Cys Glu Arg Ser Val Leu Phe Asp Met Phe His Phe Gly Phe
50 55 60

Val Ser Leu Phe Leu His Glu Trp Ile Gln Ala Lys His Phe Trp Ala
65 70 75 80

Gly Leu Phe Ile Val Leu Pro Ser Asp Val Phe Phe Ser Val His His
85 90 95

Leu Glu Ala Pro Asp Gly Ser Phe Pro Asn Ile Ala Lys Leu Ser Leu
100 105 110

Ile Ile Leu Leu Arg Xaa
115

<210> 164
<211> 43
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (43)
<223> Xaa equals stop translation

<400> 164
Met Leu Leu Gln Phe Thr Leu Trp Val Phe Gly Ala Ile His Phe Pro
1 5 10 15

Lys Cys Leu Gly Ile Lys Glu Glu Leu Leu Lys Cys Cys Leu Gln Leu
20 25 30

Pro Pro Ser Ser Thr Tyr Glu Lys Val Val Xaa
35 40

<210> 165
<211> 48
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (48)
<223> Xaa equals stop translation

<400> 165
Met Leu Ser Arg Arg Leu His Cys Leu Val Leu Tyr Phe Leu Leu Leu
1 5 10 15

Leu Leu Ser Phe Ile His Thr Leu Ser Val Ser His Ile Cys Ser Ser
20 25 30

Phe Ile Trp Leu Phe Pro Lys Asn Ile Glu Ser Glu Ala Thr Met Xaa
35 40 45

<210> 166
<211> 46
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (46)
<223> Xaa equals stop translation

<400> 166
Met Glu Lys Met Gly Gln Gly Leu Leu Ser Ser Thr Tyr Leu Thr Val
1 5 10 15

Leu His Leu Ile Gln Leu Val Gly Cys Gly Leu Leu Thr Glu Glu Ile
20 25 30

Lys Glu Ser Lys Tyr Leu Ile Lys Thr Leu Gly Ser Gly Xaa
35 40 45

<210> 167
<211> 207
<212> PRT
<213> Homo sapiens

<400> 167
Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn Cys Ile Phe Phe
1 5 10 15

Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile Thr Ala Ile Ser
20 25 30

116

Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile Phe Phe Pro Leu
35 40 45

Pro Ala Cys Leu Asn Pro Val Leu Tyr Val Phe Phe Asn Pro Lys Phe
50 55 60

Lys Glu Asp Trp Lys Leu Leu Lys Arg Arg Val Thr Lys Lys Ser Gly
65 70 75 80

Ser Val Ser Val Ser Ile Ser Ser Gln Gly Gly Cys Leu Glu Gln Asp
85 90 95

Phe Tyr Tyr Asp Cys Gly Met Tyr Ser His Leu Gln Gly Asn Leu Thr
100 105 110

Val Cys Asp Cys Cys Glu Ser Phe Leu Leu Thr Lys Pro Val Ser Cys
115 120 125

Lys His Leu Ile Lys Ser His Ser Cys Pro Ala Leu Ala Val Ala Ser
130 135 140

Cys Gln Arg Pro Glu Gly Tyr Trp Ser Asp Cys Gly Thr Gln Ser Ala
145 150 155 160

His Ser Asp Tyr Ala Asp Glu Asp Ser Phe Val Ser Asp Ser Ser
165 170 175

Asp Gln Val Gln Ala Cys Gly Arg Ala Cys Phe Tyr Gln Ser Arg Gly
180 185 190

Phe Pro Leu Val Arg Tyr Ala Tyr Asn Leu Pro Arg Val Lys Asp
195 200 205

<210> 168
<211> 51
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (51)
<223> Xaa equals stop translation

<400> 168
Met Tyr Ile Phe Glu Leu Ser Leu Tyr Leu Glu Gly Thr Ser Phe Val
1 5 10 15

Val Val Leu Leu Phe Leu Leu Ile Ser Val Ser Leu Asp Ser Pro Pro
20 25 30

Thr Thr Lys Gly Trp Asp Ser Val Leu His Ile Trp Val Pro Leu Ile
35 40 45

Val Gln Xaa
50

<210> 169
<211> 43
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (43)
<223> Xaa equals stop translation

<400> 169
Met Ala His Pro Gly Leu Pro Lys Thr Val Pro Val Tyr Ala Val Val
1 5 10 15

Leu Ala Leu Leu Ile Met Thr Leu Pro Leu Thr Leu Thr Ile Asn Leu
20 25 30

Asp Asp Asn Leu Tyr Gly Asn Ser Ala Lys Xaa
35 40

<210> 170
<211> 56
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (56)
<223> Xaa equals stop translation

<400> 170
Met Arg Pro Trp Trp Ser Leu Leu Leu Glu Ala Cys Ala Thr Cys Ala
1 5 10 15

Gln Thr Gly Pro Thr Arg Ser Thr Ser Cys Thr Gln Glu Val Ser His
20 25 30

Ser Ser Ser Thr Ala Tyr Pro Ala Pro Met Arg Arg Cys Cys Leu
35 40 45

Pro Ser Pro Arg Ser Cys Thr Xaa
50 55

<210> 171
<211> 109
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (109)
<223> Xaa equals stop translation

<400> 171
Met Ala Leu Ala Gly Ser Val Phe Val Leu Gly Gly Val Leu Val Leu

		118	
1	5	10	15
Cys Val Glu Arg Asn Gly Glu Gly Glu Met Gly Trp Pro Gln His Leu			
20		25	30
Pro Lys Ser Gln Pro Leu Ser Pro Pro Val Ala Val Arg Arg Cys Ser			
35		40	45
Phe Glu Arg Ser Trp Ile Asp Leu Leu Val Glu Thr Ser Ser Ser Met			
50		55	60
Val Thr Cys Arg Gln Gln Val Gly Thr Pro Asn Gly Met Glu Gly Arg			
65		70	75
Gly Gly Gly Pro Lys Thr Thr Phe Pro Ile Arg Leu Gln Leu Ser Gly			
85		90	95
Ala Cys Ala Val Arg Pro Glu Ile Gln Trp Glu Val Xaa			
100		105	

<210> 172
 <211> 51
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (17)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (51)
 <223> Xaa equals stop translation

<400> 172			
Met Phe Leu Phe Phe Tyr Leu Ser Leu Ala Val Tyr Ala Gln Arg Gln			
1	5	10	15
Xaa Ser Gly Ser Cys Arg Gln Thr Asp His Arg Trp Lys Ser Arg Gly			
20		25	30
Ala Arg Arg Cys Phe Leu Glu Pro Arg Asp Pro Gly Ser Val Pro Gly			
35		40	45
His Pro Xaa			
50			

<210> 173
 <211> 566
 <212> PRT
 <213> Homo sapiens

<400> 173
 Met Ala Pro Leu Ala Leu His Leu Leu Val Leu Val Pro Ile Leu Leu

1	5	10	15
---	---	----	----

119

Ser Leu Val Ala Ser Gln Asp Trp Lys Ala Glu Arg Ser Gln Asp Pro
 20 25 30

Phe Glu Lys Cys Met Gln Asp Pro Asp Tyr Glu Gln Leu Leu Lys Val
 35 40 45

Val Thr Trp Gly Leu Asn Arg Thr Leu Lys Pro Gln Arg Val Ile Val
 50 55 60

Val Gly Ala Gly Val Ala Gly Leu Val Ala Ala Lys Val Leu Ser Asp
 65 70 80

Ala Gly His Lys Val Thr Ile Leu Glu Ala Asp Asn Arg Ile Gly Gly
 85 90 95

Arg Ile Phe Thr Tyr Arg Asp Gln Asn Thr Gly Trp Ile Gly Glu Leu
 100 105 110

Gly Ala Met Arg Met Pro Ser Ser His Arg Ile Leu His Lys Leu Cys
 115 120 125

Gln Gly Leu Gly Leu Asn Leu Thr Lys Phe Thr Gln Tyr Asp Lys Asn
 130 135 140

Thr Trp Thr Glu Val His Glu Val Lys Leu Arg Asn Tyr Val Val Glu
 145 150 160

Lys Val Pro Glu Lys Leu Gly Tyr Ala Leu Arg Pro Gln Glu Lys Gly
 165 170 175

His Ser Pro Glu Asp Ile Tyr Gln Met Ala Leu Asn Gln Ala Leu Lys
 180 185 190

Asp Leu Lys Ala Leu Gly Cys Arg Lys Ala Met Lys Lys Phe Glu Arg
 195 200 205

His Thr Leu Leu Glu Tyr Leu Leu Gly Glu Gly Asn Leu Ser Arg Pro
 210 215 220

Ala Val Gln Leu Leu Gly Asp Val Met Ser Glu Asp Gly Phe Phe Tyr
 225 230 240

Leu Ser Phe Ala Glu Ala Leu Arg Ala His Ser Cys Leu Ser Asp Arg
 245 250 255

Leu Gln Tyr Ser Arg Ile Val Gly Gly Trp Asp Leu Leu Pro Arg Ala
 260 265 270

Leu Leu Ser Ser Leu Ser Gly Leu Val Leu Leu Asn Ala Pro Val Val
 275 280 285

Ala Met Thr Gln Gly Pro His Asp Val His Val Gln Ile Glu Thr Ser
 290 295 300

Pro Pro Ala Arg Asn Leu Lys Val Leu Lys Ala Asp Val Val Leu Leu
 305 310 315 320

120

Thr Ala Ser Gly Pro Ala Val Lys Arg Ile Thr Phe Ser Pro Pro Leu
325 330 335

Pro Arg His Met Gln Glu Ala Leu Arg Arg Leu His Tyr Val Pro Ala
340 345 350

Thr Lys Val Phe Leu Ser Phe Arg Arg Pro Phe Trp Arg Glu Glu His
355 360 365

Ile Glu Gly Gly His Ser Asn Thr Asp Arg Pro Ser Arg Met Ile Phe
370 375 380

Tyr Pro Pro Pro Arg Glu Gly Ala Leu Leu Ala Ser Tyr Thr Trp
385 390 395 400

Ser Asp Ala Ala Ala Ala Phe Ala Gly Leu Ser Arg Glu Glu Ala Leu
405 410 415

Arg Leu Ala Leu Asp Asp Val Ala Ala Leu His Gly Pro Val Val Arg
420 425 430

Gln Leu Trp Asp Gly Thr Gly Val Val Lys Arg Trp Ala Glu Asp Gln
435 440 445

His Ser Gln Gly Gly Phe Val Val Gln Pro Pro Ala Leu Trp Gln Thr
450 455 460

Glu Lys Asp Asp Trp Thr Val Pro Tyr Gly Arg Ile Tyr Phe Ala Gly
465 470 475 480

Glu His Thr Ala Tyr Pro His Gly Trp Val Glu Thr Ala Val Lys Leu
485 490 495

Leu Arg Ala Ala Ile Lys Ile Asn Ser Arg Lys Gly Pro Ala Ser Asp
500 505 510

Thr Ala Ser Pro Glu Gly His Ala Ser Asp Met Glu Gly Gln Gly His
515 520 525

Val His Gly Val Ala Ser Ser Pro Ser His Asp Leu Ala Lys Glu Glu
530 535 540

Gly Ser His Pro Pro Val Gln Gly Gln Leu Ser Leu Gln Asn Thr Thr
545 550 555 560

His Thr Arg Thr Ser His
565

<210> 174

<211> 224

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (76)

<223> Xaa equals any of the naturally occurring L-amino acids

121

<400> 174
 Met Ala Arg Ala Arg Gly Ser Pro Cys Pro Pro Leu Pro Pro Gly Arg
 1 5 10 15
 Met Ser Trp Pro His Gly Ala Leu Leu Phe Leu Trp Leu Phe Ser Pro
 20 25 30
 Pro Leu Gly Ala Gly Gly Gly Val Ala Val Thr Ser Ala Ala Gly
 35 40 45
 Gly Gly Ser Pro Pro Ala Thr Ser Cys Pro Val Ala Cys Ser Cys Ser
 50 55 60
 Asn Gln Ala Ser Arg Val Ile Cys Thr Arg Arg Xaa Leu Ala Glu Val
 65 70 75 80
 Pro Ala Ser Ile Pro Val Asn Thr Arg Tyr Leu Asn Leu Gln Glu Asn
 85 90 95
 Gly Ile Gln Val Ile Arg Thr Asp Thr Phe Lys His Leu Arg His Leu
 100 105 110
 Glu Ile Leu Gln Leu Ser Lys Asn Leu Val Arg Lys Ile Glu Val Gly
 115 120 125
 Ala Phe Asn Gly Leu Pro Ser Leu Asn Thr Leu Glu Leu Phe Asp Asn
 130 135 140
 Arg Leu Thr Thr Val Pro Thr Gln Ala Phe Glu Tyr Leu Ser Lys Leu
 145 150 155 160
 Arg Glu Leu Trp Leu Arg Asn Asn Pro Ile Glu Ser Ile Pro Ser Tyr
 165 170 175
 Ala Phe Asn Arg Val Pro Ser Leu Arg Arg Leu Asp Leu Gly Glu Leu
 180 185 190
 Lys Arg Leu Glu Tyr Ile Ser Glu Ala Ala Phe Glu Gly Leu Val Asn
 195 200 205
 Leu Arg Tyr Leu Asn Leu Gly Met Cys Asn Leu Lys Asp Ile Pro Asn
 210 215 220

<210> 175
<211> 123
<212> PRT
<213> Homo sapiens

<400> 175
 Met His Asp Gly Ser Lys Pro Phe Pro Arg Tyr Gly Tyr Lys Pro Ser
 1 5 10 15
 Pro Pro Asn Gly Cys Gly Ser Pro Leu Phe Gly Val His Leu Asn Ile

		122
20	25	30
Gly Ile Pro Ser Leu Thr Lys Cys Cys Asn Gln His Asp Arg Cys Tyr		
35	40	45
Glu Thr Cys Gly Lys Ser Lys Asn Asp Cys Asp Glu Glu Phe Gln Tyr		
50	55	60
Cys Leu Ser Lys Ile Cys Arg Asp Val Gln Lys Thr Leu Gly Leu Thr		
65	70	75
Gln His Val Gln Ala Cys Glu Thr Thr Val Glu Leu Leu Phe Asp Ser		
85	90	95
Val Ile His Leu Gly Cys Lys Pro Tyr Leu Asp Ser Gln Arg Ala Ala		
100	105	110
Cys Arg Cys His Tyr Glu Glu Lys Thr Asp Leu		
115	120	
<210> 176		
<211> 60		
<212> PRT		
<213> Homo sapiens		
<220>		
<221> SITE		
<222> (60)		
<223> Xaa equals stop translation		
<400> 176		
Met Gly Leu Ser Val Leu Leu Pro Leu Cys Leu Leu Gly Pro Gly Arg		
1	5	10
Phe Thr Ser Gly Gln Lys Pro Leu Asp Thr Pro Gly Leu Gly Ala Ala		
20	25	30
Val Leu Ser Val Arg Lys Ala Gly Leu Lys Met Arg Ser His Leu Thr		
35	40	45
Pro Ser Val Cys Thr Val Pro Ser Pro Gly Ser Xaa		
50	55	60
<210> 177		
<211> 105		
<212> PRT		
<213> Homo sapiens		
<400> 177		
Met Asp Thr Val Phe Leu Ile Gln Tyr Leu Phe Leu Thr Phe Pro Arg		
1	5	10
Ile Val Phe Met Leu Gly Phe Val Val Val Leu Ser Phe Leu Leu Gly		
20	25	30
Gly Tyr Leu Leu Phe Val Leu Tyr Leu Ala Ala Thr Asn Gln Thr Thr		

123

35

40

45

Asn Glu Trp Tyr Arg Gly Asp Trp Ala Trp Cys Gln Arg Cys Pro Leu
50 55 60

Val Ala Trp Pro Pro Ser Ala Glu Pro Gln Val His Arg Asn Ile His
65 70 75 80

Ser His Gly Leu Arg Ser Asn Leu Gln Glu Ile Phe Leu Pro Ala Phe
85 90 95

Pro Cys His Glu Arg Lys Lys Gln Glu
100 105

<210> 178
<211> 88
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (88)
<223> Xaa equals stop translation

<400> 178
Met Ala Asp Pro His Val Ser Phe Leu Ser Phe Arg Gln Leu Phe Ser
1 5 10 15

Trp Ala Ala Val Ile Leu Leu Arg Gly Ile Leu Gly Thr Val Ala Pro
20 25 30

Pro Pro Cys Pro Cys Val Leu Asp Leu Ala Val Tyr Pro Leu His Leu
35 40 45

Pro Val Glu Ala Pro Cys Leu Glu Val Val Phe Lys Gln Lys Asn Gly
50 55 60

Lys Asp Asn Cys Leu Val Phe Tyr Pro Asp Pro Ile Pro Leu Arg Gly
65 70 75 80

Ser Leu Leu Gly Pro Phe Ile Xaa
85.

<210> 179
<211> 88
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (55)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (66)

124

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

222 (88)

<223> Xaa equals stop translation

<400> 179

Met Ala Asp Pro His Val Ser Phe Leu Ser Phe Arg Gln Leu Phe Ser
1 5 10 15

Trp Ala Ala Val Ile Leu Leu Arg Gly Ile Leu Gly Thr Val Ala Pro
20 25 30

Pro Pro Cys Pro Cys Val Leu Asp Leu Ala Val Tyr Pro Leu His Leu
 35 40 45

Pro	Val	Glu	Ala	Pro	Cys	Xaa	Glu	Val	Val	Phe	Lys	Gln	Lys	Asn	Gly
50						55					60				

Lys	Xaa	Asn	Cys	Leu	Val	Phe	Tyr	Pro	Asp	Pro	Ile	Pro	Leu	Arg	Gly
65					70						75				80

Ser Leu Leu Gly Pro Phe Ile Xaa
85

<210> 180

<211> 49

<212> PRT

<213> Homo sapiens

<400> 180

Met	Asn	Leu	Leu	Gly	Met	Ile	Phe	Ser	Met	Cys	Gly	Leu	Met	Leu	Lys
1				5					10					15	

Leu Lys Trp Cys Ala Trp Val Ala Val Tyr Cys Ser Phe Ile Ser Phe
20 25 30

Ala Asn Ser Arg Ser Ser Glu Asp Thr Lys Gln Met Met Ser Ser Phe
35 40 45

Met

<210> 181

<211> 23

<212> PRT

<213> Homo sapiens

<400> 181

Leu Gly Ser Leu Ser Thr Ala Pro Ser Ser Ala Leu Pro Thr Leu Gly
 1 5 10 15

Ala Arg Arg Thr Arg Ser Lys
20

125

<210> 182
<211> 104
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (104)
<223> Xaa equals stop translation

<400> 182
Met Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe
1 5 10 15

Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile
20 25 30

Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe
35 40 45

Ser Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp
50 55 60

Val Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn
65 70 75 80

Tyr Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg
85 90 95

Thr Arg Val Leu Phe Ile Tyr Xaa
100

<210> 183
<211> 198
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (29)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 183
Met Lys Lys Ser Leu Glu Asn Leu Asn Arg Leu Gln Val Met Leu Leu
1 5 10 15

His Leu Thr Ala Ala Phe Leu Gln Arg Ala Gln His Xaa Phe Asp Tyr
20 25 30

Lys Asp Glu Ser Gly Phe Pro Lys Pro Pro Ser Tyr Asn Val Ala Thr
35 40 45

Thr Leu Pro Ser Tyr Asp Glu Ala Glu Arg Thr Lys Ala Glu Ala Thr
50 55 60

Ile Pro Leu Val Pro Gly Arg Asp Glu Asp Phe Val Gly Arg Asp Asp

126

65

70

75

80

Phe Asp Asp Ala Asp Gln Leu Arg Ile Gly Asn Asp Gly Ile Phe Met
 85 90 95

Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe Leu
 100 105 110

Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile Ser
 115 120 125

Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe Ser
 130 135 140

Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp Val
 145 150 155 160

Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn Tyr
 165 170 175

Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg Thr
 180 185 190

Arg Val Leu Phe Ile Tyr
 195

<210> 184

<211> 70

<212> PRT

<213> Homo sapiens

<400> 184

Met Leu Leu His Leu Thr Ala Ala Phe Leu Gln Arg Ala Gln Phe Ser
 1 5 10 15

Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp Val
 20 25 30

Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn Tyr
 35 40 45

Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg Thr
 50 55 60

Arg Val Leu Phe Ile Tyr
 65 70

<210> 185

<211> 82

<212> PRT

<213> Homo sapiens

<400> 185

Met Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe
 1 5 10 15

127

Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile
20 25 30

Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe
35 40 45

Ser Thr Tyr Phe Pro Ala Phe Met Asn Ser Leu Ser Arg Ser Lys Arg
50 55 60

Thr Pro Ala Gly Ser Glu Ser Arg Cys Arg Thr Gln Arg Asn Asn His
65 70 75 80

Leu Leu

<210> 186

<211> 45

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (28)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 186

Met Lys Lys Ser Leu Glu Asn Leu Asn Arg Leu Gln Val Met Leu Leu
1 5 10 15

His Leu Thr Ala Ala Phe Leu Gln Arg Ala His Xaa Ile Leu Thr Thr
20 25 30

Arg Met Ser Leu Gly Phe Gln Ser Pro His Leu Thr Met
35 40 45

<210> 187

<211> 34

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (34)

<223> Xaa equals stop translation

<400> 187

Met Thr Val Met Asp Pro Lys Gln Met Asn Val Ala Ala Ala Val Trp
1 5 10 15

Ala Val Val Ser Tyr Val Val Ala Asp Met Glu Glu Met Leu Pro Arg
20 25 30

Ser Xaa

128

<210> 188
<211> 232
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (232)
<223> Xaa equals stop translation

<400> 188
Met Ala Thr Leu Trp Gly Gly Leu Leu Arg Leu Gly Ser Leu Leu Ser
1 5 10 15

Leu Ser Cys Leu Ala Leu Ser Val Leu Leu Leu Ala His Cys Gln Thr
20 25 30

Pro Pro Arg Ile Ser Arg Met Ser Asp Val Asn Val Ser Ala Leu Pro
35 40 45

Ile Lys Lys Asn Ser Gly His Ile Tyr Asn Lys Asn Ile Ser Gln Lys
50 55 60

Asp Cys Asp Cys Leu His Val Val Glu Pro Met Pro Val Arg Gly Pro
65 70 75 80

Asp Val Glu Ala Tyr Cys Leu Arg Cys Glu Cys Lys Tyr Glu Glu Arg
85 90 95

Ser Ser Val Thr Ile Lys Val Thr Ile Ile Ile Tyr Leu Ser Ile Leu
100 105 110

Gly Leu Leu Leu Leu Tyr Met Val Tyr Leu Thr Leu Val Glu Pro Ile
115 120 125

Leu Lys Arg Arg Leu Phe Gly His Ala Gln Leu Ile Gln Ser Asp Asp
130 135 140

Asp Ile Gly Asp His Gln Pro Phe Ala Asn Ala His Asp Val Leu Ala
145 150 155 160

Arg Ser Arg Ser Arg Ala Asn Val Leu Asn Lys Val Glu Tyr Gly Thr
165 170 175

Ala Ala Leu Glu Ala Ser Ser Pro Arg Ala Ala Lys Ser Leu Ser Leu
180 185 190

Thr Gly Met Leu Ser Ser Ala Asn Trp Gly Ile Glu Phe Lys Val Thr
195 200 205

Arg Lys Lys Gln Ala Asp Asn Trp Lys Gly Thr Asp Trp Val Leu Leu
210 215 220

Gly Phe Ile Leu Ile Pro Cys Xaa
225 230

<210> 189

129

<211> 457
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (457)
<223> Xaa equals stop translation

<400> 189

Met Ala Ala Ala Gly Arg Leu Pro Ser Ser Trp Ala Leu Phe Ser Pro
1 5 10 15

Leu Leu Ala Gly Leu Ala Leu Gly Val Gly Pro Val Pro Ala Arg
20 25 30

Ala Leu His Asn Val Thr Ala Glu Leu Phe Gly Ala Glu Ala Trp Gly
35 40 45

Thr Leu Ala Ala Phe Gly Asp Leu Asn Ser Asp Lys Gln Thr Asp Leu
50 55 60

Phe Val Leu Arg Glu Arg Asn Asp Leu Ile Val Phe Leu Ala Asp Gln
65 70 75 80

Asn Ala Pro Tyr Phe Lys Pro Lys Val Lys Val Ser Phe Lys Asn His
85 90 95

Ser Ala Leu Ile Thr Ser Val Val Pro Gly Asp Tyr Asp Gly Asp Ser
100 105 110

Gln Met Asp Val Leu Leu Thr Tyr Leu Pro Lys Asn Tyr Ala Lys Ser
115 120 125

Glu Leu Gly Ala Val Ile Phe Trp Gly Gln Asn Gln Thr Leu Asp Pro
130 135 140

Asn Asn Met Thr Ile Leu Asn Arg Thr Phe Gln Asp Glu Pro Leu Ile
145 150 155 160

Met Asp Phe Asn Gly Asp Leu Ile Pro Asp Ile Phe Gly Ile Thr Asn
165 170 175

Glu Ser Asn Gln Pro Gln Ile Leu Leu Gly Gly Asn Leu Ser Trp His
180 185 190

Pro Ala Leu Thr Thr Ser Lys Met Arg Ile Pro His Ser His Ala
195 200 205

Phe Ile Asp Leu Thr Glu Asp Phe Thr Ala Asp Leu Phe Leu Thr Thr
210 215 220

Leu Asn Ala Thr Thr Ser Thr Phe Gln Phe Glu Ile Trp Glu Asn Leu
225 230 235 240

Asp Gly Asn Phe Ser Val Ser Thr Ile Leu Glu Lys Pro Gln Asn Met
245 250 255

130

Met Val Val Gly Gln Ser Ala Phe Ala Asp Phe Asp Gly Asp Gly His
 260 265 270

Met Asp His Leu Leu Pro Gly Cys Glu Asp Lys Asn Cys Gln Lys Ser
 275 280 285

Thr Ile Tyr Leu Val Arg Ser Gly Met Lys Gln Trp Val Pro Val Leu
 290 295 300

Gln Asp Phe Ser Asn Lys Gly Thr Leu Trp Gly Phe Val Pro Phe Val
 305 310 315 320

Asp Glu Gln Gln Pro Thr Glu Ile Pro Ile Pro Ile Thr Leu His Ile
 325 330 335

Gly Asp Tyr Asn Met Asp Gly Tyr Pro Asp Ala Leu Val Ile Leu Lys
 340 345 350

Asn Thr Ser Gly Ser Asn Gln Gln Ala Phe Leu Leu Glu Asn Val Pro
 355 360 365

Cys Asn Asn Ala Ser Cys Glu Glu Ala Arg Arg Met Phe Lys Val Tyr
 370 375 380

Trp Glu Leu Thr Asp Leu Asn Gln Ile Lys Asp Ala Met Val Ala Thr
 385 390 395 400

Phe Phe Asp Ile Tyr Glu Asp Gly Ile Leu Asp Ile Val Val Leu Ser
 405 410 415

Lys Gly Tyr Thr Lys Asn Asp Phe Ala Ile His Thr Leu Lys Asn Asn
 420 425 430

Phe Glu Ala Asp Ala Tyr Phe Val Lys Val Ile Val Leu Ser Gly Leu
 435 440 445

Cys Ser Asn Asp Cys Pro Arg Arg Xaa
 450 455

<210> 190

<211> 185

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (185)

<223> Xaa equals stop translation

<400> 190

Met Leu Phe Leu Phe Ser Met Ala Thr Leu Leu Arg Thr Ser Phe Ser
 1 5 10 15

Asp Pro Gly Val Ile Pro Arg Ala Leu Pro Asp Glu Ala Ala Phe Ile
 20 25 30

Glu Met Glu Ile Glu Ala Thr Asn Gly Ala Val Pro Gln Gly Gln Arg

131

35

40

45

Pro Pro Pro Arg Ile Lys Asn Phe Gln Ile Asn Asn Gln Ile Val Lys
 50 55 60

Leu Lys Tyr Cys Tyr Thr Cys Lys Ile Phe Arg Pro Pro Arg Ala Ser
 65 70 75 80

His Cys Ser Ile Cys Asp Asn Cys Val Glu Arg Phe Asp His His Cys
 85 90 95

Pro Trp Val Gly Asn Cys Val Gly Lys Arg Asn Tyr Arg Tyr Phe Tyr
 100 105 110

Leu Phe Ile Leu Ser Leu Ser Leu Leu Thr Ile Tyr Val Phe Ala Phe
 115 120 125

Asn Ile Val Tyr Val Ala Leu Lys Ser Leu Lys Ile Gly Phe Leu Glu
 130 135 140

Thr Leu Lys Gly Asn Ser Trp Asn Cys Ser Arg Ser Pro His Leu Leu
 145 150 155 160

Leu Tyr Thr Leu Val Arg Arg Gly Thr Asp Trp Ile Ser Tyr Phe Pro
 165 170 175

Arg Gly Ser Gln Pro Asp Asn Gln Xaa
 180 185

<210> 191

<211> 147

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (147)

<223> Xaa equals stop translation

<400> 191

Met Arg Val Leu Val Val Thr Ile Ala Pro Ile Tyr Trp Ala Leu Ala
 1 5 10 15

Arg Glu Ser Gly Glu Ala Leu Asn Gly His Ser Leu Thr Gly Lys
 20 25 30

Phe Arg Gln Ser His Thr Trp Ser Leu Leu Gln Gly Ala Ala His Asp
 35 40 45

Asp Pro Val Ala Arg Gly Leu Asp Pro Asp Gly Leu Leu Leu Asp
 50 55 60

Val Val Val Asn Gly Val Val Pro Gly Arg Ala Trp Leu Thr Gln Ile
 65 70 75 80

Phe Lys Cys Arg Thr Leu Lys Lys His Tyr Val Gln Thr Arg Ala Trp
 85 90 95

132

Pro Ala Val Arg Gly Leu His Thr Ala Leu Leu Pro Gly Arg Pro Pro
100 105 110

Leu Val Pro Thr Leu Gln Pro Gln His Pro Val Gln Arg Gly Pro Gly
115 120 125

Pro Pro Ala Pro Ala Gly Ala Ala Pro Ala Gly Leu Ser Tyr Gln Leu
130 135 140

Gly Leu Xaa
145

<210> 192

<211> 125

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (125)

<223> Xaa equals stop translation

<400> 192

Met Gly Glu Pro Asn Arg His Pro Ser Met Phe Leu Leu Leu Val
1 5 10 15

Leu Glu Arg Leu Tyr Ala Ser Pro Met Asp Gly Thr Ser Ser Ala Leu
20 25 30

Ser Met Gly Pro Phe Val Pro Phe Ile Met Arg Cys Gly His Ser Pro
35 40 45

Val Tyr His Ser Arg Glu Met Ala Ala Arg Ala Leu Val Pro Phe Val
50 55 60

Met Ile Asp His Ile Pro Asn Thr Ile Arg Thr Leu Leu Ser Thr Leu
65 70 75 80

Pro Ser Cys Thr Asp Gln Cys Phe Arg Ala Lys Pro His Ser Trp Gly
85 90 95

His Phe Ser Arg Phe His Leu Leu Gln Ala Tyr Ser Asp Ser Lys
100 105 110

Thr Arg Asn Glu Phe Arg Leu Pro Ala Arg Ala Asp Xaa
115 120 125

<210> 193

<211> 52

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (52)

133

<223> Xaa equals stop translation

<400> 193
Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn Cys Ile Phe Phe
1 5 10 15

Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile Thr Ala Ile Ser
20 25 30

Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile Phe Phe Pro Cys
35 40 45

Leu Leu Ala Xaa
50

<210> 194
<211> 320
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (68)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (115)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (213)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (320)
<223> Xaa equals stop translation

<400> 194
Met Ala Pro Leu Ala. Leu His Leu Leu Val Leu Val Pro Ile Leu Leu
1 5 10 15

Ser Leu Val Ala Ser Gln Asp Trp Lys Ala Glu Arg Ser Gln Asp Pro
20 25 30

Phe Glu Lys Cys Met Gln Asp Pro Asp Tyr Glu Gln Leu Leu Lys Val
35 40 45

Thr Ile Leu Glu Ala Asp Asn Arg Ile Gly Gly Arg Ile Phe Thr Tyr
50 55 60

Arg Asp Gln Xaa Thr Gly Trp Ile Gly Glu Leu Gly Ala Met Arg Met
65 70 75 80

Pro Ser Ser His Arg Ile Leu His Lys Leu Cys Gln Gly Leu Gly Leu

	134	
85		90
		95
Asn Leu Thr Lys Phe Thr Gln Tyr Asp Lys Asn Thr Trp Thr Glu Val		
100	105	110
His Glu Xaa Lys Leu Arg Asn Tyr Val Val Glu Lys Val Pro Glu Lys		
115	120	125
Leu Gly Tyr Ala Leu Arg Pro Gln Glu Lys Gly His Ser Pro Glu Asp		
130	135	140
Ile Tyr Gln Met Ala Leu Asn Gln Ala Leu Lys Asp Leu Lys Ala Leu		
145	150	155
160		
Gly Cys Arg Lys Ala Met Lys Lys Phe Glu Arg His Thr Leu Leu Glu		
165	170	175
Tyr Leu Leu Gly Glu Gly Asn Leu Ser Arg Pro Ala Val Gln Leu Leu		
180	185	190
Gly Asp Val Met Ser Glu Asp Gly Phe Phe Tyr Leu Ser Phe Ala Glu		
195	200	205
Ala Leu Arg Ala Xaa Ser Cys Leu Ser Asp Arg Leu Gln Tyr Ser Arg		
210	215	220
Ile Val Gly Gly Trp Asp Leu Leu Pro Arg Ala Leu Leu Ser Ser Leu		
225	230	235
240		
Ser Gly Leu Val Leu Leu Asn Ala Pro Val Val Ala Met Thr Gln Gly		
245	250	255
Pro His Asp Val His Val Gln Ile Glu Thr Ser Pro Pro Ala Arg Asn		
260	265	270
Leu Lys Val Leu Lys Ala Asp Val Val Leu Leu Thr Ala Ser Gly Pro		
275	280	285
Ala Val Lys Arg Ile Thr Phe Ser Pro Arg Cys Pro Ala Thr Cys Arg		
290	295	300
Arg Arg Cys Gly Cys Thr Thr Cys Arg Pro Pro Arg Cys Ser Xaa		
305	310	315
320		

<210> 195
<211> 130
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (38)
<223> Xaa equals any of the naturally occurring L-amino acids

135

<220>

<221> SITE

<222> (53)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 195

Pro	Phe	Cys	Ser	Gly	Phe	Phe	Pro	Ser	Leu	Trp	Ile	Tyr	Leu	Pro	Phe
1					5				10					15	

Ile	Phe	Asn	Val	Ser	Asp	Leu	Trp	Met	Gly	Ser	Leu	Ser	Gly	Cys	Ala
						20		25					30		

Leu	Pro	Phe	Cys	Leu	Xaa	Val	Phe	Phe	Leu	Thr	Val	Ser	Pro	Ser	Ala
						35		40				45			

Val	Gly	Leu	Leu	Xaa	Phe	Ala	Gly	Gly	Pro	Leu	Gln	Thr	Leu	Phe	Ala
					50		55			60					

Trp	Val	Ser	Pro	Val	Glu	Ala	Ala	Glu	Gln	Gln	Arg	Leu	Leu	Pro	Val
					65		70		75		80				

Leu	Ser	Ser	Gly	Ser	Phe	Val	Ser	Glu	Gly	Thr	Cys	Gln	Met	Pro	Ala
					85			90				95			

Arg	Ala	Leu	Leu	Tyr	Glu	Val	Ser	Val	Gly	Pro	Tyr	Trp	Glu	Ile	Pro
					100			105			110				

Pro	Ser	Gln	Asp	Thr	Arg	Arg	Ser	Gly	Thr	Tyr	Leu	Arg	Arg	Gln	Ser
					115			120			125				

Asp	Pro														
	130														

<210> 196

<211> 108

<212> PRT

<213> Homo sapiens

<400> 196

His	Glu	Gly	Ser	Cys	Arg	Ala	Pro	Gly	Phe	Ser	Ala	His	Lys	Gly	Arg
1					5				10			15			

Gly	Cys	Pro	Ser	Pro	Arg	Met	Thr	Leu	Pro	Ser	Arg	Ala	Leu	Ala	Ser
					20			25			30				

Leu	Gly	Val	Gly	Val	Trp	Gly	Met	Leu	Arg	Leu	Asn	Gln	Val	Thr	Val
					35		40		45						

Ser	Cys	Gly	Gly	Ser	Arg	Trp	Ser	Ser	Arg	Val	Ala	Leu	Gly	Ala	Phe
					50		55		60						

Ser	Trp	Val	Cys	Gly	Val	Ala	Leu	Val	Leu	Gln	Pro	Ser	Gly	Gly	Gly
					65		70		75		80				

Leu	Gly	Leu	Thr	Ser	Pro	Ser	Glu	Gly	Cys	Trp	Glu	Gly	Glu	Leu	Ala
					85		90			95					

136

Leu Ala Val Leu Arg Ala Pro Gly Gly Ser Pro Ser
100 105

<210> 197
<211> 104
<212> PRT
<213> Homo sapiens

<400> 197
Ile Pro Leu Thr Leu Pro Gly Ile Phe Leu Leu Ile Arg Leu Phe Trp
1 5 10 15

Arg Leu Gly Gln Ser Ile Cys Gly Pro Gly Lys Leu Val Leu Trp Pro
20 25 30

Gln Phe Cys Cys Gly Cys Ala Val Ile Ser Gly His Cys Val Pro Arg
35 40 45

Gly Met Pro Ser Ser Trp Leu Pro Gly Cys Phe Val Leu Leu Cys Leu
50 55 60

Val Ala Val Gly Cys Gln Leu Arg Glu Trp Gly Val Gly Val Ser
65 70 75 80

Ala Val Gly Leu Leu Ala Leu Pro His Leu Gln Val Leu Gly Met Arg
85 90 95

Gly Arg Gly Leu Ile Ser Gly Gly
100

<210> 198
<211> 237
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (142)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 198
Gly Pro Ala Gly Lys Glu Ala Trp Ile Trp Ser Trp Leu Leu Pro Ser
1 5 10 15

Pro Gly Pro Ala Pro Leu Pro Ser Ala Ser Trp Gly Leu Cys Gly Asp
20 25 30

Ala Pro Arg Ala Ala Ala Arg Gly Pro Val Glu Pro Gly Ala Ala Arg
35 40 45

Met Ala Leu Leu Ser Arg Pro Ala Leu Thr Leu Leu Leu Leu Met
50 55 60

Ala Ala Val Val Arg Cys Gln Glu Gln Ala Gln Thr Thr Asp Trp Arg
65 70 75 80

137

Ala Thr Leu Lys Thr Ile Arg Asn Gly Val His Lys Ile Asp Thr Tyr
85 90 95

Leu Asn Ala Ala Leu Asp Leu Leu Gly Gly Glu Asp Gly Leu Cys Gln
100 105 110

Tyr Lys Cys Ser Asp Gly Ser Lys Pro Phe Pro Arg Tyr Gly Tyr Lys
115 120 125

Pro Ser Pro Pro Asn Gly Cys Gly Ser Pro Leu Phe Gly Xaa His Leu
130 135 140

Asn Ile Gly Ile Pro Ser Leu Thr Lys Cys Cys Asn Gln His Asp Arg
145 150 155 160

Cys Tyr Glu Thr Cys Gly Lys Ser Lys Asn Asp Cys Asp Glu Glu Phe
165 170 175

Gln Tyr Cys Leu Ser Lys Ile Cys Arg Asp Val Gln Lys Thr Leu Gly
180 185 190

Leu Thr Gln His Val Gln Ala Cys Glu Thr Thr Val Glu Leu Leu Phe
195 200 205

Asp Ser Val Ile His Leu Gly Cys Lys Pro Tyr Leu Asp Ser Gln Arg
210 215 220

Ala Ala Cys Arg Cys His Tyr Glu Glu Lys Thr Asp Leu
225 230 235

<210> 199

<211> 8

<212> PRT

<213> Homo sapiens

<400> 199

Cys Cys Asn Gln His Asp Arg Cys
1 5

<210> 200

<211> 15

<212> PRT

<213> Homo sapiens

<400> 200

Ser Leu Thr Lys Cys Cys Asn Gln His Asp Arg Cys Tyr Glu Thr
1 5 10 15

<210> 201

<211> 16

<212> PRT

<213> Homo sapiens

<400> 201

Leu Thr Lys Cys Cys Asn Gln His Asp Arg Cys Tyr Glu Thr Cys Gly

138

1	5	10	15
---	---	----	----

<210> 202
<211> 260
<212> PRT
<213> Homo sapiens

<400> 202
Gly Thr Ser Ser Ala Arg Pro Arg Gly Ala Leu Pro Gly Gly Ser Ala
1 5 10 15

Pro Ser Ala Pro His Gly Gln Leu Pro Gly Arg Ala Gln Pro Ala Pro
20 25 30

Val Ser Gly Pro Pro Pro Thr Ser Gly Leu Cys His Phe Asp Pro Ala
35 40 45

Ala Pro Trp Pro Leu Trp Pro Gly Pro Trp Gln Leu Pro Pro His Pro
50 55 60

Gln Asp Trp Pro Ala His Pro Asp Ile Pro Gln Asp Trp Val Ser Phe
65 70 75 80

Leu Arg Ser Phe Gly Gln Leu Thr Leu Cys Pro Arg Asn Gly Thr Val
85 90 95

Thr Gly Lys Trp Arg Gly Ser His Val Val Gly Leu Leu Thr Thr Leu
100 105 110

Asn Phe Gly Asp Gly Pro Asp Arg Asn Lys Thr Arg Thr Phe Gln Ala
115 120 125

Thr Val Leu Gly Ser Gln Met Gly Leu Lys Gly Ser Ser Ala Gly Gln
130 135 140

Leu Val Leu Ile Thr Ala Arg Val Thr Thr Glu Arg Thr Ala Gly Thr
145 150 155 160

Cys Leu Tyr Phe Ser Ala Val Pro Gly Ile Leu Pro Ser Ser Gln Pro
165 170 175

Pro Ile Ser Cys Ser Glu Glu Gly Ala Gly Asn Ala Thr Leu Ser Pro
180 185 190

Arg Met Gly Glu Glu Cys Val Ser Val Trp Ser His Glu Gly Leu Val
195 200 205

Leu Thr Lys Leu Leu Thr Ser Glu Glu Leu Ala Leu Cys Gly Ser Arg
210 215 220

Leu Leu Val Leu Gly Ser Phe Leu Leu Leu Phe Cys Gly Leu Leu Cys
225 230 235 240

Cys Val Thr Ala Met Cys Phe His Pro Arg Arg Glu Ser His Trp Ser

139
245 250 255

Arg Thr Arg Leu
260

<210> 203
<211> 80
<212> PRT
<213> Homo sapiens

<400> 203
Ala Arg Ala Pro Pro Gly Pro Glu Gly Leu Ser Pro Glu Ala Gln Pro
1 5 10 15

Pro Leu Leu Pro Met Gly Asn Cys Gln Ala Gly His Asn Leu His Leu
20 25 30

Cys Leu Ala His His Pro Pro Leu Val Cys Ala Thr Leu Ile Leu Leu
35 40 45

Leu Leu Gly Leu Ser Gly Leu Gly Ser Phe Leu Leu Thr His
50 55 60

Arg Thr Gly Leu Arg Thr Leu Thr Ser Pro Arg Thr Gly Ser Leu Phe
65 70 75 80

<210> 204
<211> 224
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (6)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (9)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (22)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (143)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE

140

<222> (186)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 204

Arg Phe Leu Ser Val Xaa Pro Gln Xaa Glu Val Pro Phe Leu Leu His
1 5 10 15Pro Cys Val Cys Phe Xaa Gly Gly His Pro Ser Leu Leu Pro Asp Pro
20 25 30Cys Arg Ala Val Gly Gly Trp Glu Ala Pro Arg Cys Cys Leu His
35 40 45Glu Ala Leu Cys Gln Ser Leu Gly Cys Lys Ala Glu Glu Ile Val Ser
50 55 60Val Ser Glu Ser Ser Ala Gln Arg Cys Trp Tyr Leu Leu Arg Gly
65 70 75 80Arg Lys Ala Gly Gly Arg Gly Pro Ala Ser Pro Val Leu Phe Ala Leu
85 90 95Met Arg Leu Glu Ser Leu Cys His Leu Cys Leu Ala Cys Leu Phe Phe
100 105 110Arg Leu Pro Ala Thr Arg Thr Val Tyr Cys Met Asn Glu Ala Glu Ile
115 120 125Val Asp Val Ala Leu Gly Ile Leu Ile Glu Ser Arg Lys Gln Xaa Lys
130 135 140Ala Cys Glu Gln Pro Ala Leu Ala Gly Ala Asp Asn Pro Glu His Ser
145 150 155 160Pro Pro Cys Ser Val Ser Pro His Thr Ser Ser Gly Ser Ser Ser Glu
165 170 175Glu Glu Asp Ser Gly Lys Gln Ala Leu Xaa Pro Gly Leu Ser Pro Ser
180 185 190Gln Arg Pro Gly Gly Ser Ser Ser Ala Cys Ser Arg Ser Pro Glu Glu
195 200 205Glu Glu Glu Asp Val Leu Lys Tyr Val Arg Glu Ile Phe Phe Ser
210 215 220

<210> 205

<211> 199

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (35)

141

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (103)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (191)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 205

Val Pro Gly Trp Pro Arg Ala Cys Ser Pro Cys Gln Ala Asp Ser Pro
1 5 10 15Arg Ala His Pro Pro Lys Leu Arg Gly Ile Leu Arg Trp Ala Pro Val
20 25 30Pro Leu Xaa Cys Ala Ala Leu Cys Pro Pro Leu Asp Ser Gly Met Ser
35 40 45Met Ala Ala Cys Pro Glu Ala Pro Glu Pro Ser Phe Leu Arg Glu Val
50 55 60Pro Ser Ser Pro Ala Ser Thr Gln Trp His Arg Pro Cys Asn Phe Arg
65 70 75 80Gln Val Glu Ala Asn Pro Arg Lys Glu Pro Lys Asn Leu Val Trp Arg
85 90 95Asp Val Ser Leu Gly Gln Xaa Ser Arg Thr Pro Arg Gly Ser Gly Leu
100 105 110Glu Leu Val Arg Val Cys Gly Gly Met Gln Arg Asp Lys Thr Val
115 120 125Val Glu Glu Arg Val Gly Glu Glu Arg Glu Arg Glu Arg Glu
130 135 140Ser Leu Gly Gly Ala Gly Lys His Gly Glu Met Arg Cys Val Tyr Val
145 150 155 160Arg Glu Ser Val Gly Ala Pro Gly Arg Ala Gly Gly Gly Asn Gly
165 170 175Val Asn Ser Val Gly Cys Val Arg Thr Val His Ser Gly Ser Xaa Pro
180 185 190Pro Pro Ser Ala Gly Val Ser
195

<210> 206

<211> 174

<212> PRT

<213> Homo sapiens

142

<400> 206

Thr	Arg	Pro	Gly	Lys	Glu	Leu	Asn	Leu	Val	Phe	Gly	Leu	Gln	Leu	Ser
1					5				10				15		

Met	Ala	Arg	Ile	Gly	Ser	Thr	Val	Asn	Met	Asn	Leu	Met	Gly	Trp	Leu
						20		25				30			

Tyr	Ser	Lys	Ile	Glu	Ala	Leu	Leu	Gly	Ser	Ala	Gly	His	Thr	Thr	Leu
						35		40			45				

Gly	Ile	Thr	Leu	Met	Ile	Gly	Gly	Ile	Thr	Cys	Ile	Leu	Ser	Leu	Ile
						50		55			60				

Cys	Ala	Leu	Ala	Leu	Ala	Tyr	Leu	Asp	Gln	Arg	Ala	Glu	Arg	Ile	Leu
						65		70		75		80			

His	Lys	Glu	Gln	Gly	Lys	Thr	Gly	Glu	Val	Ile	Lys	Leu	Thr	Asp	Val
						85		90		95					

Lys	Asp	Phe	Ser	Leu	Pro	Leu	Trp	Leu	Ile	Phe	Ile	Ile	Cys	Val	Cys
						100		105		110					

Tyr	Tyr	Val	Ala	Val	Phe	Pro	Phe	Ile	Gly	Leu	Gly	Lys	Val	Phe	Phe
						115		120		125					

Thr	Glu	Lys	Phe	Gly	Phe	Ser	Ser	Gln	Ala	Ala	Ser	Ala	Ile	Asn	Ser
						130		135		140					

Val	Val	Tyr	Val	Ile	Ser	Ala	Pro	Met	Ser	Pro	Val	Phe	Gly	Leu	Leu
						145		150		155		160			

Val	Asp	Lys	Thr	Gly	Lys	Asn	Ile	Ile	Trp	Val	Leu	Cys	Ala		
						165		170							

<210> 207

<211> 31

<212> PRT

<213> Homo sapiens

<400> 207

Cys	Lys	Asp	Leu	Cys	Ser	Arg	Val	Tyr	Leu	Leu	Thr	Leu	Ser	Pro	Leu
1					5				10		15				

Leu	Ser	Tyr	Asp	Pro	Ala	Thr	Ser	His	Ser	Pro	Arg	Asn	Thr	Gln	
						20		25		30					

<210> 208

<211> 369

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (78)

<223> Xaa equals any of the naturally occurring L-amino acids

143

<400> 208
 Ile Ile Cys Glu Cys Trp Glu Glu Glu Cys Gln Ser Cys Arg Leu Lys
 1 5 10 15

Ile Thr Gln Pro Arg Glu Ile Cys Arg Met Asp Phe Leu Val Leu Phe
 20 25 30

Leu Phe Tyr Leu Ala Ser Val Leu Met Gly Leu Val Leu Ile Cys Val
 35 40 45

Cys Ser Lys Thr His Ser Leu Lys Gly Leu Ala Arg Gly Gly Ala Gln
 50 55 60

Ile Phe Ser Cys Ile Ile Pro Glu Cys Leu Gln Arg Ala Xaa His Gly
 65 70 75 80

Leu Leu His Tyr Leu Phe His Thr Arg Asn His Thr Phe Ile Val Leu
 85 90 95

His Leu Val Leu Gln Gly Met Val Tyr Thr Glu Tyr Thr Trp Glu Val
 100 105 110

Phe Gly Tyr Cys Gln Glu Leu Glu Leu Ser Leu His Tyr Leu Leu Leu
 115 120 125

Pro Tyr Leu Leu Leu Gly Val Asn Leu Phe Phe Phe Thr Leu Thr Cys
 130 135 140

Gly Thr Asn Pro Gly Ile Ile Thr Lys Ala Asn Glu Leu Leu Phe Leu
 145 150 155 160

His Val Tyr Glu Phe Asp Glu Val Met Phe Pro Lys Asn Val Arg Cys
 165 170 175

Ser Thr Cys Asp Leu Arg Lys Pro Ala Arg Ser Lys His Cys Ser Val
 180 185 190

Cys Asn Trp Cys Val His Arg Phe Asp His His Cys Val Trp Val Asn
 195 200 205

Asn Cys Ile Gly Ala Trp Asn Ile Arg Tyr Phe Leu Ile Tyr Val Leu
 210 215 220

Thr Leu Thr Ala Ser Ala Ala Thr Val Ala Ile Val Ser Thr Thr Phe
 225 230 235 240

Leu Val His Leu Val Val Met Ser Asp Leu Tyr Gln Glu Thr Tyr Ile
 245 250 255

Asp Asp Leu Gly His Leu His Val Met Asp Thr Val Phe Leu Ile Gln
 260 265 270

Tyr Leu Phe Leu Thr Phe Pro Arg Ile Val Phe Met Leu Gly Phe Val
 275 280 285

Val Val Leu Ser Phe Leu Leu Gly Gly Tyr Leu Leu Phe Val Leu Tyr
 290 295 300

144

Leu Ala Ala Thr Asn Gln Thr Thr Asn Glu Trp Tyr Arg Gly Asp Trp
 305 310 315 320

Ala Trp Cys Gln Arg Cys Pro Leu Val Ala Trp Pro Pro Ser Ala Glu
 325 330 335

Pro Gln Val His Arg Asn Ile His Ser His Gly Leu Arg Ser Asn Leu
 340 345 350

Gln Glu Ile Phe Leu Pro Ala Phe Pro Cys His Glu Arg Lys Lys Gln
 355 360 365

Glu

<210> 209

<211> 147

<212> PRT

<213> Homo sapiens

<400> 209

Leu Leu Ser Phe Lys Ile Arg Gly Leu Arg Thr Glu Asp Ala Gly Trp
 1 5 10 15

Ala Gln Ser Ser Ser Gly Gly Leu Cys Val Arg Gly Asp Ala Phe Trp
 20 25 30

Met Pro Ser Ser Ser Gly Leu Gly Ser Pro Ser Arg Pro Pro Ser
 35 40 45

Ser Phe Leu Cys Leu Leu Leu Leu Leu Pro Pro Ala Ala Leu Ala
 50 55 60

Leu Leu Leu Phe Phe Leu Asp Phe Phe Pro Pro Arg Ala Ala Val Ser
 65 70 75 80

Pro Phe Leu Pro Asp His Cys Ser Ala Arg Gln Pro Arg Val Trp Arg
 85 90 95

Arg Glu Thr Leu Asn Arg Ser Ala Ser Gly Leu Gly Cys Trp Ala Arg
 100 105 110

Ser Thr Glu Gln Gly Ala Val Gly Val Ala Thr Gly Thr Val Leu Asp
 115 120 125

Ile Ser Leu Pro Ala Ser Cys Leu Ser Leu Trp Pro Pro Gly Pro Ser
 130 135 140

Gly Gly Ile
 145

<210> 210

<211> 143

<212> PRT

<213> Homo sapiens

145

<400> 210
 Gln Leu Gly Leu Cys Leu Thr Ser Ala Ser Leu Pro Pro Ala Ser Arg
 1 5 10 15
 Cys Gly His Gln Ala Pro Leu Gly Ala Ser Asp Leu Ser Ala His His
 20 25 30
 Ser Ala Pro Gly Phe Ser Asp Ser Tyr Phe Thr Met Ser Cys Gln Ser
 35 40 45
 Ser Leu Ser Arg Ala Glu Ile Leu Gln Cys Pro Leu Val Pro Ser Val
 50 55 60
 Ser Pro Pro Thr His Leu Pro Gln Gly Arg Ala Asn Lys Ser Ser Arg
 65 70 75 80
 Ala Ser Leu Pro Leu Leu Pro Gln Thr His Trp Cys Leu Phe Pro Ser
 85 90 95
 Ala Arg Gly Trp Arg Arg Gly Ile Gln Ser Gly Leu Pro Pro Gly Gly
 100 105 110
 Ser Cys Thr Ser Pro Arg Ser Pro Pro Gln Thr Leu His Gln His Ile
 115 120 125
 Thr Leu Val Asn His Asn Thr Ser Tyr Trp Gln Ser Pro Ser Thr
 130 135 140

<210> 211
 <211> 160
 <212> PRT
 <213> Homo sapiens

<400> 211
 His Gln Pro Pro Cys Leu Leu Pro Leu Ala Val Ala Thr Arg Pro Leu
 1 5 10 15
 Trp Gly His Leu Thr Cys Leu Pro Ile Ile Leu His Leu Val Ser Val
 20 25 30
 Thr Leu Thr Ser Pro Cys Leu Ala Asn Gln Ala Phe Gln Gly Gln Arg
 35 40 45
 Ser Tyr Asn Ala Leu Trp Cys Pro Leu Phe Leu Leu Leu Pro Thr Ser
 50 55 60
 Pro Lys Gly Glu Gln Thr Asn His Pro Glu Pro Ala Cys Pro Cys Phe
 65 70 75 80
 Pro Lys Leu Thr Gly Val Phe Ser Leu Gln His Val Val Gly Ala Glu
 85 90 95
 Glu Phe Ser Gln Val Phe Leu Leu Val Asp Pro Val Pro Val Leu Asp
 100 105 110
 His Leu Leu Lys Leu Phe Thr Ser Thr Ser His Leu Leu Ile Ile Ile
 115 120 125

146

Pro His Ile Gly Lys Ala Pro Ala Pro Asp Ser Leu Leu Glu Glu Leu			
130	135	140	
Ser Leu Ser Leu Ala Thr His Cys Lys Val Ala Val Ala Arg Phe Thr			
145	150	155	160

<210> 212
<211> 157
<212> PRT
<213> Homo sapiens

<400> 212
Met Ala Ala Glu Gly Ser Arg Phe Ser Ser Gln Ser Pro Gly Leu Val
1 5 10 15

Asp Arg Gln Gly Pro Lys Cys Asp Pro Ser Arg Leu Val Ser Pro Trp
20 25 30

Gly Arg His Gly Leu Arg Ile Leu Gln Ile Gly His His His Gly Arg
35 40 45

Asp Gly Gln His Glu Ala Thr His His Leu Leu Arg Val Leu Arg Ala
50 55 60

Pro Arg Val Gly Lys Ala Asp Glu Gly Ala Val Asp Ser Asp Pro Ser
65 70 75 80

Thr Pro Leu Gln Leu Lys His Glu Ala Ala His Ala Glu Asp His Ala
85 90 95

Gln Gln Val His Val Val Arg Arg Val Val Gln Gly Arg Val Thr
100 105 110

Phe Ala Arg Arg Gly Leu Val Pro Gln His Phe Val Arg Pro Pro Trp
115 120 125

Val Arg His Ile Val Ser Gly His Ser Glu Ser Lys Ala Arg Ser Arg
130 135 140

Leu Phe Arg Cys Arg Asn Arg Ser Phe Arg Arg Ala Ser
145 150 155

<210> 213
<211> 38
<212> PRT
<213> Homo sapiens

<400> 213
Arg Leu Val Ser Pro Trp Gly Arg His Gly Leu Arg Ile Leu Gln Ile
1 5 10 15

Gly His His His Gly Arg Asp Gly Gln His Glu Ala Thr His His Leu

147

20

25

30

Leu Arg Val Leu Arg Ala
35

<210> 214
<211> 12
<212> PRT
<213> Homo sapiens

<400> 214
Pro Thr Asp Val Leu Lys Ile Arg Met Gln Ala Gln
1 5 10

<210> 215
<211> 7
<212> PRT
<213> Homo sapiens

<400> 215
Thr Tyr Glu Gln Leu Lys Arg
1 5

<210> 216
<211> 137
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (22)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (33)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (71)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 216
Arg Pro Arg Pro Ser Ala Ser Ser Leu Ala Arg Ser Ala Ser Leu Leu
1 5 10 15

Pro Ala Ala His Gly Xaa Gly Val Gly Gly Ala Gly Gly Ser Ser
20 25 30

Xaa Leu Arg Ser Arg Tyr Gln Gln Leu Gln Asn Glu Glu Glu Ser Gly
35 40 45

Glu Pro Glu Gln Ala Ala Gly Asp Ala Pro Pro Pro Tyr Ser Ser Ile
50 55 60

148

Ser Ala Glu Ser Ala His Xaa Phe Asp Tyr Lys Asp Glu Ser Gly Phe
65 70 75 80

Pro Lys Pro Pro Ser Tyr Asn Val Ala Thr Thr Leu Pro Ser Tyr Asp
85 90 95

Glu Ala Glu Arg Thr Lys Ala Glu Ala Thr Ile Pro Leu Val Pro Gly
100 105 110

Arg Asp Glu Asp Phe Val Gly Arg Asp Asp Phe Asp Asp Ala Asp Gln
115 120 125

Leu Arg Ile Gly Asn Asp Gly Ile Phe
130 135

<210> 217

<211> 20

<212> PRT

<213> Homo sapiens

<400> 217

Arg Tyr Gln Gln Leu Gln Asn Glu Glu Ser Gly Glu Pro Glu Gln
1 5 10 15

Ala Ala Gly Asp
20

<210> 218

<211> 22

<212> PRT

<213> Homo sapiens

<400> 218

Pro Gly Arg Asp Glu Asp Phe Val Gly Arg Asp Asp Phe Asp Asp Ala
1 5 10 15

Asp Gln Leu Arg Ile Gly
20

<210> 219

<211> 103

<212> PRT

<213> Homo sapiens

<400> 219

Met Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe
1 5 10 15

Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile
20 25 30

Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe
35 40 45

149

Ser Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp
50 55 60
Val Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn
65 70 75 80
Tyr Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg
85 90 95
Thr Arg Val Leu Phe Ile Tyr
100

<210> 220
<211> 198
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (29)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 220
Met Lys Lys Ser Leu Glu Asn Leu Asn Arg Leu Gln Val Met Leu Leu
1 5 10 15

His Leu Thr Ala Ala Phe Leu Gln Arg Ala Gln His Xaa Phe Asp Tyr
20 25 30

Lys Asp Glu Ser Gly Phe Pro Lys Pro Pro Ser Tyr Asn Val Ala Thr
35 40 45

Thr Leu Pro Ser Tyr Asp Glu Ala Glu Arg Thr Lys Ala Glu Ala Thr
50 55 60

Ile Pro Leu Val Pro Gly Arg Asp Glu Asp Phe Val Gly Arg Asp Asp
65 70 75 80

Phe Asp Asp Ala Asp Gln Leu Arg Ile Gly Asn Asp Gly Ile Phe Met
85 90 95

Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe Leu
100 105 110

Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile Ser
115 120 125

Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe Ser
130 135 140

Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp Val
145 150 155 160

Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn Tyr
165 170 175

Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg Thr

150
180 185 190

Arg Val Leu Phe Ile Tyr
195

<210> 221
<211> 70
<212> PRT
<213> Homo sapiens

<400> 221
Met Leu Leu His Leu Thr Ala Ala Phe Leu Gln Arg Ala Gln Phe Ser
1 5 10 15

Thr Tyr Phe Pro Gly Tyr Phe Asp Gly Gln Tyr Trp Leu Trp Trp Val
20 25 30

Phe Leu Val Leu Gly Phe Leu Leu Phe Leu Arg Gly Phe Ile Asn Tyr
35 40 45

Ala Lys Val Arg Lys Met Pro Glu Thr Phe Ser Asn Leu Pro Arg Thr
50 55 60

Arg Val Leu Phe Ile Tyr
65 70

<210> 222
<211> 82
<212> PRT
<213> Homo sapiens

<400> 222
Met Leu Thr Phe Phe Met Ala Phe Leu Phe Asn Trp Ile Gly Phe Phe
1 5 10 15

Leu Ser Phe Cys Leu Thr Thr Ser Ala Ala Gly Arg Tyr Gly Ala Ile
20 25 30

Ser Gly Phe Gly Leu Ser Leu Ile Lys Trp Ile Leu Ile Val Arg Phe
35 40 45

Ser Thr Tyr Phe Pro Ala Phe Met Asn Ser Leu Ser Arg Ser Lys Arg
50 55 60

Thr Pro Ala Gly Ser Glu Ser Arg Cys Arg Thr Gln Arg Asn Asn His
65 70 75 80

Leu Leu

<210> 223
<211> 45
<212> PRT
<213> Homo sapiens

151

<220>

<221> SITE

<222> (28)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 223

Met	Lys	Lys	Ser	Leu	Glu	Asn	Leu	Asn	Arg	Leu	Gln	Val	Met	Leu	Leu
1				5				10					15		

His	Leu	Thr	Ala	Ala	Phe	Leu	Gln	Arg	Ala	His	Xaa	Ile	Leu	Thr	Thr
						20				25			30		

Arg	Met	Ser	Leu	Gly	Phe	Gln	Ser	Pro	His	Leu	Thr	Met			
						35			40			45			

<210> 224

<211> 33

<212> PRT

<213> Homo sapiens

<400> 224

Met	Thr	Val	Met	Asp	Pro	Lys	Gln	Met	Asn	Val	Ala	Ala	Val	Trp	
1					5				10				15		

Ala	Val	Val	Ser	Tyr	Val	Val	Ala	Asp	Met	Glu	Glu	Met	Leu	Pro	Arg
					20				25			30			

Ser

<210> 225

<211> 189

<212> PRT

<213> Homo sapiens

<400> 225

Pro	Arg	Val	Arg	Ser	Arg	Glu	Pro	Val	Ala	Gly	Ala	Pro	Gly	Cys	Gly
1						5			10			15			

Thr	Ala	Gly	Pro	Pro	Ala	Met	Ala	Thr	Leu	Trp	Gly	Gly	Leu	Leu	Arg
						20			25			30			

Leu	Gly	Ser	Leu	Leu	Ser	Leu	Ser	Cys	Leu	Ala	Leu	Ser	Val	Leu	Leu
						35			40			45			

Leu	Ala	His	Cys	Gln	Thr	Pro	Pro	Ser	Asp	Cys	Leu	His	Val	Val	Glu
					50			55			60				

Pro	Met	Pro	Val	Arg	Gly	Pro	Asp	Val	Glu	Ala	Tyr	Cys	Leu	Arg	Cys
65						70			75			80			

Glu	Cys	Lys	Tyr	Glu	Glu	Arg	Ser	Ser	Val	Thr	Ile	Lys	Val	Thr	Ile
						85			90			95			

Ile	Ile	Tyr	Leu	Ser	Ile	Leu	Gly	Leu	Leu	Leu	Tyr	Met	Val	Tyr	
						100			105			110			

152

Leu Thr Leu Val Glu Pro Ile Leu Lys Arg Arg Leu Phe Gly His Ala
115 120 125

Gln Leu Ile Gln Ser Asp Asp Asp Ile Gly Asp His Gln Pro Phe Ala
130 135 140

Asn Ala His Asp Val Leu Ala Arg Ser Arg Ser Arg Ala Asn Val Leu
145 150 155 160

Asn Lys Val Glu Tyr Ala Gln Gln Arg Trp Lys Leu Gln Val Gln Glu
165 170 175

Gln Arg Lys Ser Val Phe Asp Arg His Val Val Leu Ser
180 185

<210> 226
<211> 231
<212> PRT
<213> Homo sapiens

<400> 226
Met Ala Thr Leu Trp Gly Gly Leu Leu Arg Leu Gly Ser Leu Leu Ser
1 5 10 15

Leu Ser Cys Leu Ala Leu Ser Val Leu Leu Leu Ala His Cys Gln Thr
20 25 30

Pro Pro Arg Ile Ser Arg Met Ser Asp Val Asn Val Ser Ala Leu Pro
35 40 45

Ile Lys Lys Asn Ser Gly His Ile Tyr Asn Lys Asn Ile Ser Gln Lys
50 55 60

Asp Cys Asp Cys Leu His Val Val Glu Pro Met Pro Val Arg Gly Pro
65 70 75 80

Asp Val Glu Ala Tyr Cys Leu Arg Cys Glu Cys Lys Tyr Glu Glu Arg
85 90 95

Ser Ser Val Thr Ile Lys Val Thr Ile Ile Tyr Leu Ser Ile Leu
100 105 110

Gly Leu Leu Leu Tyr Met Val Tyr Leu Thr Leu Val Glu Pro Ile
115 120 125

Leu Lys Arg Arg Leu Phe Gly His Ala Gln Leu Ile Gln Ser Asp Asp
130 135 140

Asp Ile Gly Asp His Gln Pro Phe Ala Asn Ala His Asp Val Leu Ala
145 150 155 160

Arg Ser Arg Ser Arg Ala Asn Val Leu Asn Lys Val Glu Tyr Gly Thr
165 170 175

Ala Ala Leu Glu Ala Ser Ser Pro Arg Ala Ala Lys Ser Leu Ser Leu
180 185 190

153

Thr Gly Met Leu Ser Ser Ala Asn Trp Gly Ile Glu Phe Lys Val Thr
 195 200 205
 Arg Lys Lys Gln Ala Asp Asn Trp Lys Gly Thr Asp Trp Val Leu Leu
 210 215 220
 Gly Phe Ile Leu Ile Pro Cys
 225 230

 <210> 227
 <211> 456
 <212> PRT
 <213> Homo sapiens

 <400> 227
 Met Ala Ala Ala Gly Arg Leu Pro Ser Ser Trp Ala Leu Phe Ser Pro
 1 5 10 15

 Leu Leu Ala Gly Leu Ala Leu Leu Gly Val Gly Pro Val Pro Ala Arg
 20 25 30

 Ala Leu His Asn Val Thr Ala Glu Leu Phe Gly Ala Glu Ala Trp Gly
 35 40 45

 Thr Leu Ala Ala Phe Gly Asp Leu Asn Ser Asp Lys Gln Thr Asp Leu
 50 55 60

 Phe Val Leu Arg Glu Arg Asn Asp Leu Ile Val Phe Leu Ala Asp Gln
 65 70 75 80

 Asn Ala Pro Tyr Phe Lys Pro Lys Val Lys Val Ser Phe Lys Asn His
 85 90 95

 Ser Ala Leu Ile Thr Ser Val Val Pro Gly Asp Tyr Asp Gly Asp Ser
 100 105 110

 Gln Met Asp Val Leu Leu Thr Tyr Leu Pro Lys Asn Tyr Ala Lys Ser
 115 120 125

 Glu Leu Gly Ala Val Ile Phe Trp Gly Gln Asn Gln Thr Leu Asp Pro
 130 135 140

 Asn Asn Met Thr Ile Leu Asn Arg Thr Phe Gln Asp Glu Pro Leu Ile
 145 150 155 160

 Met Asp Phe Asn Gly Asp Leu Ile Pro Asp Ile Phe Gly Ile Thr Asn
 165 170 175

 Glu Ser Asn Gln Pro Gln Ile Leu Leu Gly Gly Asn Leu Ser Trp His
 180 185 190

 Pro Ala Leu Thr Thr Ser Lys Met Arg Ile Pro His Ser His Ala
 195 200 205

 Phe Ile Asp Leu Thr Glu Asp Phe Thr Ala Asp Leu Phe Leu Thr Thr
 210 215 220

154

Leu Asn Ala Thr Thr Ser Thr Phe Gln Phe Glu Ile Trp Glu Asn Leu
 225 230 235 240
 Asp Gly Asn Phe Ser Val Ser Thr Ile Leu Glu Lys Pro Gln Asn Met
 245 250 255
 Met Val Val Gly Gln Ser Ala Phe Ala Asp Phe Asp Gly Asp Gly His
 260 265 270
 Met Asp His Leu Leu Pro Gly Cys Glu Asp Lys Asn Cys Gln Lys Ser
 275 280 285
 Thr Ile Tyr Leu Val Arg Ser Gly Met Lys Gln Trp Val Pro Val Leu
 290 295 300
 Gln Asp Phe Ser Asn Lys Gly Thr Leu Trp Gly Phe Val Pro Phe Val
 305 310 315 320
 Asp Glu Gln Gln Pro Thr Glu Ile Pro Ile Pro Ile Thr Leu His Ile
 325 330 335
 Gly Asp Tyr Asn Met Asp Gly Tyr Pro Asp Ala Leu Val Ile Leu Lys
 340 345 350
 Asn Thr Ser Gly Ser Asn Gln Gln Ala Phe Leu Leu Glu Asn Val Pro
 355 360 365
 Cys Asn Asn Ala Ser Cys Glu Glu Ala Arg Arg Met Phe Lys Val Tyr
 370 375 380
 Trp Glu Leu Thr Asp Leu Asn Gln Ile Lys Asp Ala Met Val Ala Thr
 385 390 395 400
 Phe Phe Asp Ile Tyr Glu Asp Gly Ile Leu Asp Ile Val Val Leu Ser
 405 410 415
 Lys Gly Tyr Thr Lys Asn Asp Phe Ala Ile His Thr Leu Lys Asn Asn
 420 425 430
 Phe Glu Ala Asp Ala Tyr Phe Val Lys Val Ile Val Leu Ser Gly Leu
 435 440 445
 Cys Ser Asn Asp Cys Pro Arg Arg
 450 455

 <210> 228
 <211> 282
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (144)
 <223> Xaa equals any of the naturally occurring L-amino acids
 <220>

155

<221> SITE

<222> (168)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 228

```

Met Thr Lys Arg Glu Asp Gly Gly Tyr Thr Phe Thr Ala Thr Pro Glu
   1          5          10          15

```

Asp Phe Pro Lys Lys His Lys Ala Pro Val Ile Asp Ile Gly Ile Ala
20 25 30

Asn Thr Gly Lys Phe Ile Met Thr Ala Ser Ser Asp Thr Thr Val Leu
35 40 45

Ile	Trp	Ser	Leu	Lys	Gly	Gln	Val	Leu	Ser	Thr	Ile	Asn	Thr	Asn	Gln
50						55					60				

Met Asn Asn Thr His Ala Ala Val Ser Pro Cys Gly Arg Phe Val Ala
65 70 75 80

Ser Cys Gly Phe Thr Pro Asp Val Lys Val Trp Glu Val Cys Phe Gly
85 90 95

Lys Lys Gly Glu Phe Gln Glu Val Val Arg Ala Phe Glu Leu Lys Gly
 100 105 110

His Ser Ala Ala Val His Ser Phe Ala Phe Ser Asn Asp Ser Arg Arg
115 120 125

Met Ala Ser Val Ser Lys Asp Gly Thr Trp Lys Leu Trp Asp Thr Xaa
 130 135 140

Val Glu Tyr Lys Lys Lys Gln Asp Pro Tyr Leu Leu Lys Thr Gly Arg
145 150 155 160

Phe Glu Glu Ala Ala Gly Ala Xaa Pro Cys Arg Leu Ala Leu Ser Pro
165 170 175

Asn Ala Gln Val Leu Ala Leu Ala Ser Gly Ser Ser Ile His Leu Tyr
180 185 190

Glu Cys Ile Ala Asn Leu Ser Phe Asp Ile Thr Gly Arg Phe Leu Ala
210 215 220

Ser Cys Gly Asp Arg Ala Val Arg Leu Phe His Asn Thr Pro Gly His
225 230 235 240

Arg Ala Met Val Glu Glu Met Gln Gly His Leu Lys Arg Ala Ser Asn
245 250 255

Glu Ser Thr Arg Gln Arg Leu Gln Gln Gln Leu Thr Gln Ala Gln Glu
260 265 270

Thr Leu Lys Ser Leu Gly Ala Leu Lys Lys
275 280

<210> 229
<211> 456
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (17)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (37)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (318)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (342)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 229
Val Ile Arg His Glu Gly Ser Thr Asn Met Glu Leu Ser Gln Met Ser
1 5 10 15

Xaa Leu Met Gly Leu Ser Val Leu Leu Gly Leu Leu Ala Leu Met Ala
20 25 30

Thr Ala Ala Val Xaa Arg Gly Trp Leu Arg Ala Gly Glu Arg Ser
35 40 45

Gly Arg Pro Ala Cys Gln Lys Ala Asn Gly Phe Pro Pro Asp Lys Ser
50 55 60

Ser Gly Ser Lys Lys Gln Lys Gln Tyr Gln Arg Ile Arg Lys Glu Lys
65 70 75 80

Pro Gln Gln His Asn Phe Thr His Arg Leu Leu Ala Ala Leu Lys
85 90 95

Ser His Ser Gly Asn Ile Ser Cys Met Asp Phe Ser Ser Asn Gly Lys
100 105 110

Tyr Leu Ala Thr Cys Ala Asp Asp Arg Thr Ile Arg Ile Trp Ser Thr
115 120 125

Lys Asp Phe Leu Gln Arg Glu His Arg Ser Met Arg Ala Asn Val Glu
130 135 140

Leu Asp His Ala Thr Leu Val Arg Phe Ser Pro Asp Cys Arg Ala Phe
145 150 155 160

157

Ile Val Trp Leu Ala Asn Gly Asp Thr Leu Arg Val Phe Lys Met Thr
165 170 175

Lys Arg Glu Asp Gly Gly Tyr Thr Phe Thr Ala Thr Pro Glu Asp Phe
180 185 190

Pro Lys Lys His Lys Ala Pro Val Ile Asp Ile Gly Ile Ala Asn Thr
195 200 205

Gly Lys Phe Ile Met Thr Ala Ser Ser Asp Thr Thr Val Leu Ile Trp
210 215 220

Ser Leu Lys Gly Gln Val Leu Ser Thr Ile Asn Thr Asn Gln Met Asn
225 230 235 240

Asn Thr His Ala Ala Val Ser Pro Cys Gly Arg Phe Val Ala Ser Cys
245 250 255

Gly Phe Thr Pro Asp Val Lys Val Trp Glu Val Cys Phe Gly Lys Lys
260 265 270

Gly Glu Phe Gln Glu Val Val Arg Ala Phe Glu Leu Lys Gly His Ser
275 280 285

Ala Ala Val His Ser Phe Ala Phe Ser Asn Asp Ser Arg Arg Met Ala
290 295 300

Ser Val Ser Lys Asp Gly Thr Trp Lys Leu Trp Asp Thr Xaa Val Glu
305 310 315 320

Tyr Lys Lys Gln Asp Pro Tyr Leu Leu Lys Thr Gly Arg Phe Glu
325 330 335

Glu Ala Ala Gly Ala Xaa Pro Cys Arg Leu Ala Leu Ser Pro Asn Ala
340 345 350

Gln Val Leu Ala Leu Ala Ser Gly Ser Ser Ile His Leu Tyr Asn Thr
355 360 365

Arg Arg Gly Glu Lys Glu Glu Cys Phe Glu Arg Val His Gly Glu Cys
370 375 380

Ile Ala Asn Leu Ser Phe Asp Ile Thr Gly Arg Phe Leu Ala Ser Cys
385 390 395 400

Gly Asp Arg Ala Val Arg Leu Phe His Asn Thr Pro Gly His Arg Ala
405 410 415

Met Val Glu Glu Met Gln Gly His Leu Lys Arg Ala Ser Asn Glu Ser
420 425 430

Thr Arg Gln Arg Leu Gln Gln Leu Thr Gln Ala Gln Glu Thr Leu
435 440 445

Lys Ser Leu Gly Ala Leu Lys Lys
450 455

158

<210> 230
<211> 363
<212> PRT
<213> Homo sapiens

<400> 230
Met Ser Val Met Val Val Arg Lys Lys Val Thr Arg Lys Trp Glu Lys
1 5 10 15
Leu Pro Gly Arg Asn Thr Phe Cys Cys Asp Gly Arg Val Met Met Ala
20 25 30
Arg Gln Lys Gly Ile Phe Tyr Leu Thr Leu Phe Leu Ile Leu Gly Thr
35 40 45
Cys Thr Leu Phe Phe Ala Phe Glu Cys Arg Tyr Leu Ala Val Gln Leu
50 55 60
Ser Pro Ala Ile Pro Val Phe Ala Ala Met Leu Phe Leu Phe Ser Met
65 70 75 80
Ala Thr Leu Leu Arg Thr Ser Phe Ser Asp Pro Gly Val Ile Pro Arg
85 90 95
Ala Leu Pro Asp Glu Ala Ala Phe Ile Glu Met Glu Ile Glu Ala Thr
100 105 110
Asn Gly Ala Val Pro Gln Gly Gln Arg Pro Pro Pro Arg Ile Lys Asn
115 120 125
Phe Gln Ile Asn Asn Gln Ile Val Lys Leu Lys Tyr Cys Tyr Thr Cys
130 135 140
Lys Ile Phe Arg Pro Pro Arg Ala Ser His Cys Ser Ile Cys Asp Asn
145 150 155 160
Cys Val Glu Arg Phe Asp His His Cys Pro Trp Val Gly Asn Cys Val
165 170 175
Gly Lys Arg Asn Tyr Arg Tyr Phe Tyr Leu Phe Ile Leu Ser Leu Ser
180 185 190
Leu Leu Thr Ile Tyr Val Phe Ala Phe Asn Ile Val Tyr Val Ala Leu
195 200 205
Lys Ser Leu Lys Ile Gly Phe Leu Glu Thr Leu Lys Glu Thr Pro Gly
210 215 220
Thr Val Leu Glu Val Leu Ile Cys Phe Phe Thr Leu Trp Ser Val Val
225 230 235 240
Gly Leu Thr Gly Phe His Thr Phe Leu Val Ala Leu Asn Gln Thr Thr
245 250 255
Asn Glu Asp Ile Lys Gly Ser Trp Thr Gly Lys Asn Arg Val Gln Asn
260 265 270
Pro Tyr Ser His Gly Asn Ile Val Lys Asn Cys Cys Glu Val Leu Cys

		159
275	280	285
 Gly Pro Leu Pro Pro Ser Val Leu Asp Arg Arg Gly Ile Leu Pro Leu 290 295 300		
 Glu Glu Ser Gly Ser Arg Pro Pro Ser Thr Gln Glu Thr Ser Ser Ser 305 310 315 320		
 Leu Leu Pro Gln Ser Pro Ala Pro Thr Glu Leu Asn Ser Asn Glu Met 325 330 335		
 Pro Glu Asp Ser Ser Thr Pro Glu Glu Met Pro Pro Pro Glu Pro Pro 340 345 350		
 Glu Pro Pro Gln Glu Ala Ala Glu Ala Glu Lys 355 360		
 <210> 231		
<211> 184		
<212> PRT		
<213> Homo sapiens		
 <400> 231		
Met Leu Phe Leu Phe Ser Met Ala Thr Leu Leu Arg Thr Ser Phe Ser 1 5 10 15		
 Asp Pro Gly Val Ile Pro Arg Ala Leu Pro Asp Glu Ala Ala Phe Ile 20 25 30		
 Glu Met Glu Ile Glu Ala Thr Asn Gly Ala Val Pro Gln Gly Gln Arg 35 40 45		
 Pro Pro Pro Arg Ile Lys Asn Phe Gln Ile Asn Asn Gln Ile Val Lys 50 55 60		
 Leu Lys Tyr Cys Tyr Thr Cys Lys Ile Phe Arg Pro Pro Arg Ala Ser 65 70 75 80		
 His Cys Ser Ile Cys Asp Asn Cys Val Glu Arg Phe Asp His His Cys 85 90 95		
 Pro Trp Val Gly Asn.Cys Val Gly Lys Arg Asn Tyr Arg Tyr Phe Tyr 100 105 110		
 Leu Phe Ile Leu Ser Leu Ser Leu Leu Thr Ile Tyr Val Phe Ala Phe 115 120 125		
 Asn Ile Val Tyr Val Ala Leu Lys Ser Leu Lys Ile Gly Phe Leu Glu 130 135 140		
 Thr Leu Lys Gly Asn Ser Trp Asn Cys Ser Arg Ser Pro His Leu Leu 145 150 155 160		
 Leu Tyr Thr Leu Val Arg Arg Gly Thr Asp Trp Ile Ser Tyr Phe Pro 165 170 175		
 Arg Gly Ser Gln Pro Asp Asn Gln		

160

180

<210> 232
<211> 52
<212> PRT
<213> Homo sapiens

<400> 232
Leu His Glu Cys Leu Pro Gly Ser Ile Ser Tyr Leu His Pro Arg Thr
1 5 10 15

Pro Trp Leu Cys Leu Pro Pro Gln His Leu Ser Phe Ser Thr Phe Ser
20 25 30

Pro Pro Trp Gln Pro Ala Met Ser Pro Val Pro Gly Thr Gly Gly Pro
35 40 45

Pro Cys Gly Leu
50

<210> 233
<211> 177
<212> PRT
<213> Homo sapiens

<400> 233
Met Leu Pro Leu Leu Ile Ile Cys Leu Leu Pro Ala Ile Glu Gly Lys
1 5 10 15

Asn Cys Leu Arg Cys Trp Pro Glu Leu Ser Ala Leu Ile Asp Tyr Asp
20 25 30

Leu Gln Ile Leu Trp Val Thr Pro Gly Pro Pro Thr Glu Leu Ser Gln
35 40 45

Ser Ile His Ser Leu Phe Leu Glu Asp Asn Asn Phe Leu Lys Pro Trp
50 55 60

Tyr Leu Asp Arg Asp His Leu Glu Glu Glu Thr Ala Lys Phe Phe Thr
65 70 75 80

Gln Val His Gln Ala Ile Lys Thr Leu Arg Asp Asp Lys Thr Val Leu
85 90 95

Leu Glu Glu Ile Tyr Thr His Lys Asn Leu Phe Thr Glu Arg Leu Asn
100 105 110

Lys Ile Ser Asp Gly Leu Lys Glu Lys Gly Ala Pro Pro Leu His Glu
115 120 125

Cys Leu Pro Gly Ser Ile Ser Tyr Leu His Pro Arg Thr Pro Trp Leu
130 135 140

Cys Leu Pro Pro Gln His Leu Ser Phe Ser Thr Phe Ser Pro Pro Trp
145 150 155 160

161

Gln Pro Ala Met Ser Pro Val Pro Gly Thr Gly Gly Pro Pro Cys Gly
165 170 175

Leu

<210> 234
<211> 95
<212> PRT
<213> *Homo sapiens*

<400> 234
Pro Pro Val Pro Pro Trp Ile Ser Leu Pro Leu Thr Gly Ser Pro Pro
1 5 10 15

Arg Pro Gly Phe Val Pro Val Ser Pro Phe Cys Phe Ser Pro Met Thr
 20 25 30

Asn Gly His Gln Val Leu Leu Leu Leu Leu Leu Thr Ser Ala Val Ala
35 40 45

Ala Gly Pro Trp Pro Gln Val His Ala Gly Gln Trp Gly Trp Met Cys
50 55 60

Leu Pro Pro Gly Leu Pro Ser Val Gln Ala Arg Ser Gly Leu Gly Gly
65 70 75 80

Leu Pro Gly Gly Pro Gln Trp Val Pro Gly Gly Ala Arg Gly Tyr
85 90 95

<210> 235
<211> 404
<212> PRT
<213> *Homo sapiens*

<400> 235
Ile Gln Gln Trp Gly Asp Ser Val Leu Gly Arg Arg Cys Arg Asp Leu
1 5 10 15

Leu Leu Gln Leu Tyr Leu Gln Arg Pro Glu Leu Arg Val Pro Val Pro
20 . 25 30

Glu Val Leu Leu His Ser Glu Gly Ala Ala Ser Ser Ser Val Cys Lys
 35 40 45

Leu Asp Gly Leu Ile His Arg Phe Ile Thr Leu Leu Ala Asp Thr Ser
50 55 60

Asp	Ser	Arg	Ala	Leu	Glu	Asn	Arg	Gly	Ala	Asp	Ala	Ser	Met	Ala	Cys
65					70						75				80

Arg Lys Leu Ala Val Ala His Pro Leu Leu Leu Leu Arg His Leu Pro
85 90 95

Met Ile Ala Ala Leu Leu His Gly Arg Thr His Leu Asn Phe Gln Glu
100 105 110

162

Phe Arg Gln Gln Asn His Leu Ser Cys Phe Leu His Val Leu Gly Leu
115 120 125

Leu Glu Leu Leu Gln Pro His Val Phe Arg Ser Glu His Gln Gly Ala
130 135 140

Leu Trp Asp Cys Leu Leu Ser Phe Ile Arg Leu Leu Leu Asn Tyr Arg
145 150 155 160

Lys Ser Ser Arg His Leu Ala Ala Phe Ile Asn Lys Phe Val Gln Phe
165 170 175

Ile His Lys Tyr Ile Thr Tyr Asn Ala Pro Ala Ala Ile Ser Phe Leu
180 185 190

Gln Lys His Ala Asp Pro Leu His Asp Leu Ser Phe Asp Asn Ser Asp
195 200 205

Leu Val Met Leu Lys Ser Leu Leu Ala Gly Leu Ser Leu Pro Ser Arg
210 215 220

Asp Asp Arg Thr Asp Arg Gly Leu Asp Glu Glu Gly Glu Glu Ser
225 230 235 240

Ser Ala Gly Ser Leu Pro Leu Val Ser Val Ser Leu Phe Thr Pro Leu
245 250 255

Thr Ala Ala Glu Met Ala Pro Tyr Met Lys Arg Leu Ser Arg Gly Gln
260 265 270

Thr Val Glu Asp Leu Leu Glu Val Leu Ser Asp Ile Asp Glu Met Ser
275 280 285

Arg Arg Arg Pro Glu Ile Leu Ser Phe Phe Ser Thr Asn Leu Gln Arg
290 295 300

Leu Met Ser Ser Ala Glu Glu Cys Cys Arg Asn Leu Ala Phe Ser Leu
305 310 315 320

Ala Leu Arg Ser Met Gln Asn Ser Pro Ser Ile Ala Ala Ala Phe Leu
325 330 335

Pro Thr Phe Met Tyr Cys Leu Gly Ser Gln Asp Phe Glu Val Val Gln
340 345 350

Thr Ala Leu Arg Asn Leu Pro Glu Tyr Ala Leu Leu Cys Gln Glu His
355 360 365

Ala Ala Val Leu Leu His Arg Ala Phe Leu Val Gly Met Tyr Gly Gln
370 375 380

Met Asp Pro Ser Ala Gln Ile Ser Glu Ala Leu Arg Ile Leu His Met
385 390 395 400

Glu Ala Val Met

163

<210> 236
<211> 361
<212> PRT
<213> Homo sapiens

<400> 236
Met Leu Leu Lys His Leu Gln Arg Met Val Ser Val Pro Gln Val Lys
1 5 10 15
Ala Ser Ala Leu Lys Val Val Thr Leu Thr Ala Asn Asp Lys Thr Ser
20 25 30
Val Ser Phe Ser Ser Leu Pro Gly Gln Gly Val Ile Tyr Asn Val Ile
35 40 45
Val Trp Asp Pro Phe Leu Asn Thr Ser Ala Ala Tyr Ile Pro Ala His
50 55 60
Thr Tyr Ala Cys Ser Phe Glu Ala Gly Glu Gly Ser Cys Ala Ser Leu
65 70 75 80
Gly Arg Val Ser Ser Lys Val Phe Phe Thr Leu Phe Ala Leu Leu Gly
85 90 95
Phe Phe Ile Cys Phe Phe Gly His Arg Phe Trp Lys Thr Glu Leu Phe
100 105 110
Phe Ile Gly Phe Ile Ile Met Gly Phe Phe Phe Tyr Ile Leu Ile Thr
115 120 125
Arg Leu Thr Pro Ile Lys Tyr Asp Val Asn Leu Ile Leu Thr Ala Val
130 135 140
Thr Gly Ser Val Gly Gly Met Phe Leu Val Ala Val Trp Trp Arg Phe
145 150 155 160
Gly Ile Leu Ser Ile Cys Met Leu Cys Val Gly Leu Val Leu Gly Phe
165 170 175
Leu Ile Ser Ser Val Thr Phe Phe Thr Pro Leu Gly Asn Leu Lys Ile
180 185 190
Phe His Asp Asp Gly Val Phe Trp Val Thr Phe Ser Cys Ile Ala Ile
195 200 205
Leu Ile Pro Val Val Phe Met Gly Cys Leu Arg Ile Leu Asn Ile Leu
210 215 220
Thr Cys Gly Val Ile Gly Ser Tyr Ser Val Val Leu Ala Ile Asp Ser
225 230 235 240
Tyr Trp Ser Thr Ser Leu Ser Tyr Ile Thr Leu Asn Val Leu Lys Arg
245 250 255
Ala Leu Asn Lys Asp Phe His Arg Ala Phe Thr Asn Val Pro Phe Gln
260 265 270

164

Thr Asn Asp Phe Ile Ile Leu Ala Val Trp Gly Met Leu Ala Val Ser
275 280 285

Gly Ile Thr Leu Gln Ile Arg Arg Glu Arg Gly Arg Pro Phe Phe Pro
290 295 300

Pro	His	Pro	Tyr	Lys	Leu	Trp	Lys	Gln	Glu	Arg	Glu	Arg	Arg	Val	Thr
305					310					315					320

Asn Ile Leu Asp Pro Ser Tyr His Ile Pro Pro Leu Arg Glu Arg Leu
325 330 335

Tyr Gly Arg Leu Thr Gln Ile Lys Gly Leu Phe Gln Lys Glu Gln Pro
340 345 350

Ala Gly Glu Arg Thr Pro Leu Leu Leu
355 360

<210> 237
<211> 116
<212> PRT
<213> *Homo sapiens*

<220>
<221> SITE
<222> (37)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (40)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 237
Trp Ala Arg Leu Arg Gly Pro Gly Ala His Ala Arg Thr Ser Pro Gln
1 5 10 15

Pro Trp Arg Gly Pro Ser Pro Ala Gln Ala Ala Met Gly Phe Leu Gln
20 25 30

Leu Leu Val Val Xaa Val Leu Xaa Ser Glu His Arg Val Ala Gly Ala
35 . 40 45

Ala Glu Val Phe Gly Asn Ser Ser Glu Gly Leu Ile Glu Phe Ser Val
50 55 60

Gly Lys Phe Arg Tyr Phe Glu Leu Asn Arg Pro Phe Pro Glu Glu Ala
 65 70 75 80

Ile Leu His Asp Ile Ser Ser Asn Val Thr Phe Leu Ile Phe Gln Ile
85 90 95

His Ser Gln Tyr Gln Asn Thr Thr Val Ser Phe Ser Pro Arg Arg Arg

Ser Pro Thr Met

165

<210> 238
<211> 166
<212> PRT
<213> Homo sapiens

<400> 238
Pro Arg Val Arg Pro Ala Ser Pro Pro Val Arg Ser Pro Ala Arg Trp
1 5 10 15
Gly Ser Met Ala Gly Ser Pro Leu Leu Trp Gly Pro Arg Ala Gly Gly
20 25 30
Val Gly Leu Leu Val Leu Leu Leu Gly Leu Phe Arg Pro Pro Pro
35 40 45
Ala Leu Cys Ala Arg Pro Val Lys Glu Pro Arg Gly Leu Ser Ala Ala
50 55 60
Ser Pro Pro Leu Ala Arg Leu Ala Leu Ala Ala Ser Gly Gly Gln
65 70 75 80
Cys Pro Glu Val Arg Arg Gly Arg Cys Arg Pro Gly Ala Gly Ala
85 90 95
Gly Ala Ser Ala Gly Ala Glu Arg Gln Glu Arg Ala Arg Ala Glu Ala
100 105 110
Gln Arg Leu Arg Ile Ser Arg Arg Ala Ser Trp Arg Ser Cys Cys Ala
115 120 125
Ser Gly Ala Pro Pro Ala Thr Leu Ile Arg Leu Trp Ala Trp Thr Thr
130 135 140
Thr Pro Thr Arg Leu Gln Arg Ser Ser Leu Ala Leu Cys Ser Ala Pro
145 150 155 160
Ala Leu Thr Leu Pro Pro
165

<210> 239
<211> 414
<212> PRT
<213> Homo sapiens

<400> 239
Pro Arg Val Arg Leu Ala Thr Pro Asn Ile Trp Asp Leu Ser Met Leu
1 5 10 15
Phe Ala Phe Ile Ser Leu Leu Val Met Leu Pro Thr Trp Trp Ile Val
20 25 30
Ser Ser Trp Leu Val Trp Gly Val Ile Leu Phe Val Tyr Leu Val Ile
35 40 45
Arg Ala Leu Arg Leu Trp Arg Thr Ala Lys Leu Gln Val Thr Leu Lys

166

50	55	60	
Lys Tyr Ser Val His Leu Glu Asp Met Ala Thr Asn Ser Arg Ala Phe			
65	70	75	80
Thr Asn Leu Val Arg Lys Ala Leu Arg Leu Ile Gln Glu Thr Glu Val			
85	90	95	
Ile Ser Arg Gly Phe Thr Leu Val Ser Ala Ala Cys Pro Phe Asn Lys			
100	105	110	
Ala Gly Gln His Pro Ser Gln His Leu Ile Gly Leu Arg Lys Ala Val			
115	120	125	
Tyr Arg Thr Leu Arg Ala Asn Phe Gln Ala Ala Arg Leu Ala Thr Leu			
130	135	140	
Tyr Met Leu Lys Asn Tyr Pro Leu Asn Ser Glu Ser Asp Asn Val Thr			
145	150	155	160
Asn Tyr Ile Cys Val Val Pro Phe Lys Glu Leu Gly Leu Gly Leu Ser			
165	170	175	
Glu Glu Gln Ile Ser Glu Glu Glu Ala His Asn Phe Thr Asp Gly Phe			
180	185	190	
Ser Leu Pro Ala Leu Lys Val Leu Phe Gln Leu Trp Val Ala Gln Ser			
195	200	205	
Ser Glu Phe Phe Arg Arg Leu Ala Leu Leu Ser Thr Ala Asn Ser			
210	215	220	
Pro Pro Gly Pro Leu Leu Thr Pro Ala Leu Leu Pro His Arg Ile Leu			
225	230	235	240
Ser Asp Val Thr Gln Gly Leu Pro His Ala His Ser Ala Cys Leu Glu			
245	250	255	
Glu Leu Lys Arg Ser Tyr Glu Phe Tyr Arg Tyr Phe Glu Thr Gln His			
260	265	270	
Gln Ser Val Pro Gln Cys Leu Ser Lys Thr Gln Gln Lys Ser Arg Glu			
275	280	285	
Leu Asn Asn Val His Thr Ala Val Arg Ser Leu Gln Leu His Leu Lys			
290	295	300	
Ala Leu Leu Asn Glu Val Ile Ile Leu Glu Asp Glu Leu Glu Lys Leu			
305	310	315	320
Val Cys Thr Lys Glu Thr Gln Glu Leu Val Ser Glu Ala Tyr Pro Ile			
325	330	335	
Leu Glu Gln Lys Leu Lys Leu Ile Gln Pro His Val Gln Ala Ser Asn			
340	345	350	
Asn Cys Trp Glu Glu Ala Ile Ser Gln Val Asp Lys Leu Leu Arg Arg			
355	360	365	

167

Asn Thr Asp Lys Lys Gly Lys Pro Glu Ile Ala Cys Glu Asn Pro His
370 375 380

Cys Thr Val Ser Thr Phe Glu Ala Ala Tyr Ser Thr His Cys Arg Gln
385 390 395 400

Arg Ser Asn Pro Arg Gly Ala Gly Ile Arg Ser Leu Cys Arg
405 410

<210> 240
<211> 145
<212> PRT
<213> Homo sapiens

<400> 240
Ala Ala Pro His Pro Pro Leu Leu Arg Pro Leu Cys Leu Trp Cys Pro
1 5 10 15

Leu Trp Pro Ala Trp Pro Leu Arg Gly Arg Pro Arg Ser Ala Trp Lys
20 25 30

Arg Trp Pro Pro Leu Pro Val Gly Pro Ala Lys Leu Gly Cys Ser Met
35 40 45

Thr Thr Arg Gln Pro Thr Ala Val Ser Trp Pro Cys Trp Leu Met Ser
50 55 60

Ser Ser Leu Ser Thr Ala Cys Leu Ala Trp Thr Leu Thr Gly Ser Leu
65 70 75 80

Ala Arg Glu Ala Thr Arg Arg Ala Arg Ser Leu Ser Pro Thr Trp Asn
85 90 95

Cys Ser Ala Arg Gln Val Pro Pro Ser Pro Pro His Ser Gly Leu Gly
100 105 110

Arg Arg Gly Trp Ala His Cys His Leu Thr Cys Leu Leu Val Thr Gln
115 120 125

Leu Phe Arg Val Gly Arg Ile His Pro Ile Leu Ser Leu Pro Leu Val
130 135 140

Thr
145

<210> 241
<211> 72
<212> PRT
<213> Homo sapiens

<400> 241
Leu Gln Leu Ala Ser Gln Ser Ala Gly Ile Lys Gly Met Ser His Cys
1 5 10 15

Ala Arg Pro Thr Phe Leu Thr Leu Leu Ala Ser Cys Phe Trp Ala

		168	
	20	25	30
Ala Ala Ile Pro Asn Arg Asn Val Ile Leu Ser Val Ser Phe Arg Pro			
35	40	45	
Leu His Met Gln Phe Thr Leu Ser Ile Leu Val Phe Ile Leu Arg Ile			
50	55	60	
Leu Ile Leu Leu Arg Ser Phe Leu			
65	70		
<210> 242			
<211> 140			
<212> PRT			
<213> Homo sapiens			
<400> 242			
Met Val Leu Val Leu Arg His Pro Leu Cys Ala Arg Glu Arg Ala Phe			
1	5	10	15
Arg Glu Pro Gly Arg Gly Leu Leu Thr Arg Thr Gly Gln His Asp Gly			
20	25	30	
Ala Pro Ala Val Thr Ala Val Pro Gly Pro Leu Gly Ala Val Ala Ala			
35	40	45	
Ala Glu Gly Arg Arg Ser Ala Trp Gly Ala Gly Gly Ser Ser Pro Pro			
50	55	60	
Arg Lys Val Leu Trp Gly Asp Met Arg Gly Arg Arg Ala Gly Val Asp			
65	70	75	80
Val Leu Gly Pro Ala Leu Ser Ser Glu Ala Ala Gly Ala Glu Ala Arg			
85	90	95	
Gly Trp Gly Met Pro Gly Met Gly Val Gly Val Gly Ala Ser Glu Thr			
100	105	110	
Arg Gly Ala Leu Phe Leu Gly Arg Glu Gly Val His Gly Pro Cys Pro			
115	120	125	
Met Asp Gly Leu Gly Pro Trp Pro Trp Gly Pro Trp			
130	135	140	
<210> 243			
<211> 353			
<212> PRT			
<213> Homo sapiens			
<400> 243			
Met Gly Pro Ala Val Lys Met Trp Thr Asn Ala Trp Lys Gly Leu Asp			
1	5	10	15
Asp Cys His Tyr Asn Gln Leu Cys Glu Asn Thr Pro Gly Gly His Arg			
20	25	30	

169

Cys Ser Cys Pro Arg Gly Tyr Arg Met Gln Gly Pro Ser Leu Pro Cys
 35 40 45

Leu Asp Val Asn Glu Cys Leu Gln Leu Pro Lys Ala Cys Ala Tyr Gln
 50 55 60

Cys His Asn Leu Gln Gly Ser Tyr Arg Cys Leu Cys Pro Pro Gly Gln
 65 70 75 80

Thr Leu Leu Arg Asp Gly Lys Ala Cys Thr Ser Leu Glu Arg Asn Gly
 85 90 95

Gln Asn Val Thr Thr Val Ser His Arg Gly Pro Leu Leu Pro Trp Leu
 100 105 110

Arg Pro Trp Ala Ser Ile Pro Gly Thr Ser Tyr His Ala Trp Val Ser
 115 120 125

Leu Arg Pro Gly Pro Met Ala Leu Ser Ser Val Gly Arg Ala Trp Cys
 130 135 140

Pro Pro Gly Phe Ile Arg Gln Asn Gly Val Cys Thr Asp Leu Asp Glu
 145 150 155 160

Cys Arg Val Arg Asn Leu Cys Gln His Ala Cys Arg Asn Thr Glu Gly
 165 170 175

Ser Tyr Gln Cys Leu Cys Pro Ala Gly Tyr Arg Leu Leu Pro Ser Gly
 180 185 190

Lys Asn Cys Gln Asp Ile Asn Glu Cys Glu Glu Ser Ile Glu Cys
 195 200 205

Gly Pro Gly Gln Met Cys Phe Asn Thr Arg Gly Ser Tyr Gln Cys Val
 210 215 220

Asp Thr Pro Cys Pro Ala Thr Tyr Arg Gln Gly Pro Ser Pro Gly Thr
 225 230 235 240

Cys Phe Arg Arg Cys Ser Gln Asp Cys Gly Thr Gly Pro Ser Thr
 245 250 255

Leu Gln Tyr Arg Leu Leu Pro Leu Pro Leu Gly Val Arg Ala His His
 260 265 270

Asp Val Ala Arg Leu Thr Ala Phe Ser Glu Val Gly Val Pro Ala Asn
 275 280 285

Arg Thr Glu Leu Ser Met Leu Glu Pro Asp Pro Arg Ser Pro Phe Ala
 290 295 300

Leu Arg Pro Leu Arg Ala Gly Leu Gly Ala Val Tyr Thr Arg Arg Ala
 305 310 315 320

Leu Thr Arg Ala Gly Leu Tyr Arg Leu Thr Val Arg Ala Ala Pro
 325 330 335

Arg His Gln Ser Val Phe Val Leu Leu Ile Ala Val Ser Pro Tyr Pro

170

340

345

350

Tyr

<210> 244

<211> 146

<212> PRT

<213> Homo sapiens

<400> 244

Met Arg Val Leu Val Val Thr Ile Ala Pro Ile Tyr Trp Ala Leu Ala
1 5 10 15Arg Glu Ser Gly Glu Ala Leu Asn Gly His Ser Leu Thr Gly Gly Lys
20 25 30Phe Arg Gln Ser His Thr Trp Ser Leu Leu Gln Gly Ala Ala His Asp
35 40 45Asp Pro Val Ala Arg Gly Leu Asp Pro Asp Gly Leu Leu Leu Asp
50 55 60Val Val Val Asn Gly Val Val Pro Gly Arg Ala Trp Leu Thr Gln Ile
65 70 75 80Phe Lys Cys Arg Thr Leu Lys Lys His Tyr Val Gln Thr Arg Ala Trp
85 90 95Pro Ala Val Arg Gly Leu His Thr Ala Leu Leu Pro Gly Arg Pro Pro
100 105 110Leu Val Pro Thr Leu Gln Pro Gln His Pro Val Gln Arg Gly Pro Gly
115 120 125Pro Pro Ala Pro Ala Gly Ala Ala Pro Ala Gly Leu Ser Tyr Gln Leu
130 135 140Gly Leu
145

<210> 245

<211> 638

<212> PRT

<213> Homo sapiens

<400> 245

His Ala Ser Gly Ala Phe Leu Val Val Arg Gly Glu Pro Gln Gly Ser
1 5 10 15Trp Gly Ser Met Thr Gly Val Ile Asn Gly Arg Lys Phe Gly Val Ala
20 25 30Thr Leu Asn Thr Ser Val Met Gln Glu Ala His Ser Gly Val Ser Ser
35 40 45

171

Ile His Ser Ser Ile Arg His Val Pro Ala Asn Val Gly Pro Leu Met
 50 55 60

Arg Val Leu Val Val Thr Ile Ala Pro Ile Tyr Trp Ala Leu Ala Arg
 65 70 75 80

Glu Ser Gly Glu Ala Leu Asn Gly His Ser Leu Thr Gly Gly Lys Phe
 85 90 95

Arg Gln Glu Ser His Val Glu Phe Ala Thr Gly Glu Leu Leu Thr Met
 100 105 110

Thr Gln Trp Pro Gly Val Trp Ile Pro Met Ala Ser Cys Ser Ser Thr
 115 120 125

Trp Trp Ser Met Ala Leu Ser Pro Asp Ser Leu Ala Asp Ala Asp Leu
 130 135 140

Gln Val Gln Asp Phe Glu Glu His Tyr Val Gln Thr Gly Pro Gly Gln
 145 150 155 160

Leu Phe Val Gly Ser Thr Gln Arg Phe Phe Gln Gly Gly Leu Pro Ser
 165 170 175

Phe Leu Arg Cys Asn His Ser Ile Gln Tyr Asn Ala Ala Arg Gly Pro
 180 185 190

Gln Pro Gln Leu Val Gln His Leu Arg Ala Ser Ala Ile Ser Ser Ala
 195 200 205

Phe Asp Pro Glu Ala Glu Ala Leu Arg Phe Gln Leu Ala Thr Ala Leu
 210 215 220

Gln Ala Glu Glu Asn Glu Val Gly Cys Pro Glu Gly Phe Glu Leu Asp
 225 230 235 240

Ser Gln Gly Ala Phe Cys Val Asp Val Asp Glu Cys Ala Trp Asp Ala
 245 250 255

His Leu Cys Arg Glu Gly Gln Arg Cys Val Asn Leu Leu Gly Ser Tyr
 260 265 270

Arg Cys Leu Pro Asp Cys Gly Pro Gly Phe Arg Val Ala Asp Gly Ala
 275 280 285

Gly Cys Glu Asp Val Asp Glu Cys Leu Glu Gly Leu Asp Asp Cys His
 290 295 300

Tyr Asn Gln Leu Cys Glu Asn Thr Pro Gly Gly His Arg Cys Ser Cys
 305 310 315 320

Pro Arg Gly Tyr Arg Met Gln Gly Pro Ser Leu Pro Cys Leu Asp Val
 325 330 335

Asn Glu Cys Leu Gln Leu Pro Lys Ala Cys Ala Tyr Gln Cys His Asn
 340 345 350

Leu Gln Gly Ser Tyr Arg Cys Leu Cys Pro Pro Gly Gln Thr Leu Leu

172

355

360

365

Arg Asp Gly Lys Ala Cys Thr Ser Leu Glu Arg Asn Gly Gln Asn Val
370 375 380

Thr Thr Val Ser His Arg Gly Pro Leu Leu Pro Trp Leu Arg Pro Trp
385 390 395 400

Ala Ser Ile Pro Gly Thr Ser Tyr His Ala Trp Val Ser Leu Arg Pro
405 410 415

Gly Pro Met Ala Leu Ser Ser Val Gly Arg Ala Trp Cys Pro Pro Gly
420 425 430

Phe Ile Arg Gln Asn Gly Val Cys Thr Asp Leu Asp Glu Cys Arg Val
435 440 445

Arg Asn Leu Cys Gln His Ala Cys Arg Asn Thr Glu Gly Ser Tyr Gln
450 455 460

Cys Leu Cys Pro Ala Gly Tyr Arg Leu Leu Pro Ser Gly Lys Asn Cys
465 470 475 480

Gln Asp Ile Asn Glu Cys Glu Glu Ser Ile Glu Cys Gly Pro Gly
485 490 495

Gln Met Cys Phe Asn Thr Arg Gly Ser Tyr Gln Cys Val Asp Thr Pro
500 505 510

Cys Pro Ala Thr Tyr Arg Gln Gly Pro Ser Pro Gly Thr Cys Phe Arg
515 520 525

Arg Cys Ser Gln Asp Cys Gly Thr Gly Pro Ser Thr Leu Gln Tyr
530 535 540

Arg Leu Leu Pro Leu Pro Leu Gly Val Arg Ala His His Asp Val Ala
545 550 555 560

Arg Leu Thr Ala Phe Ser Glu Val Gly Val Pro Ala Asn Arg Thr Glu
565 570 575

Leu Ser Met Leu Glu Pro Asp Pro Arg Ser Pro Phe Ala Leu Arg Pro
580 585 590

Leu Arg Ala Gly Leu Gly Ala Val Tyr Thr Arg Arg Ala Leu Thr Arg
595 600 605

Ala Gly Leu Tyr Arg Leu Thr Val Arg Ala Ala Ala Pro Arg His Gln
610 615 620

Ser Val Phe Val Leu Leu Ile Ala Val Ser Pro Tyr Pro Tyr
625 630 635

<210> 246

<211> 367

<212> PRT

<213> Homo sapiens

173

<400> 246
Met Gly Glu Lys Phe Leu Leu Leu Ala Met Lys Glu Asn His Pro Glu
1 5 10 15

Cys Phe Cys Lys Ile Leu Lys Ile Leu His Cys Met Asp Pro Gly Glu
20 25 30

Trp Leu Pro Gln Thr Glu His Cys Val His Leu Thr Pro Lys Glu Phe
35 40 45

Leu Ile Trp Thr Met Asp Ile Ala Ser Asn Glu Arg Ser Glu Ile Gln
50 55 60

Ser Val Ala Leu Arg Leu Ala Ser Lys Val Ile Ser His His Met Gln
65 70 75 80

Thr Cys Val Glu Asn Arg Glu Leu Ile Ala Ala Glu Leu Lys Gln Trp
85 90 95

Val Gln Leu Val Ile Leu Ser Cys Glu Asp His Leu Pro Thr Glu Ser
100 105 110

Arg Leu Ala Val Val Glu Val Leu Thr Ser Thr Pro Leu Phe Leu
115 120 125

Thr Asn Pro His Pro Ile Leu Glu Leu Gln Asp Thr Leu Ala Leu Trp
130 135 140

Lys Cys Val Leu Thr Leu Leu Gln Ser Glu Glu Gln Ala Val Arg Asp
145 150 155 160

Ala Ala Thr Glu Thr Val Thr Ala Met Ser Gln Glu Asn Thr Cys
165 170 175

Gln Ser Thr Glu Phe Ala Phe Cys Gln Val Asp Ala Ser Ile Ala Leu
180 185 190

Ala Leu Ala Leu Ala Val Leu Cys Asp Leu Leu Gln Gln Trp Asp Gln
195 200 205

Leu Ala Pro Gly Leu Pro Ile Leu Leu Gly Trp Leu Leu Gly Glu Ser
210 215 220

Asp Asp Leu Val Ala Cys Val Glu Ser Met His Gln Val Glu Glu Asp
225 230 235 240

Tyr Leu Phe Glu Lys Ala Glu Val Asn Phe Trp Ala Glu Thr Leu Ile
245 250 255

Phe Val Lys Tyr Leu Cys Lys His Leu Phe Cys Leu Leu Ser Lys Ser
260 265 270

Gly Trp Arg Pro Pro Ser Pro Glu Met Leu Cys His Leu Gln Arg Met
275 280 285

Val Ser Glu Gln Cys His Leu Leu Ser Gln Phe Phe Arg Glu Leu Pro
290 295 300

174

Pro Ala Ala Glu Phe Val Lys Thr Val Glu Phe Thr Arg Leu Arg Ile
305 310 315 320

Gln Glu Glu Arg Thr Leu Ala Cys Leu Arg Leu Leu Ala Phe Leu Glu
325 330 335

Gly Lys Glu Gly Glu Asp Thr Leu Val Leu Ser Val Trp Asp Ser Tyr
340 345 350

Ala Glu Ser Arg Gln Leu Thr Leu Pro Arg Thr Glu Ala Ala Cys
355 360 365

<210> 247

<211> 124

<212> PRT

<213> Homo sapiens

<400> 247

Met Gly Glu Pro Asn Arg His Pro Ser Met Phe Leu Leu Leu Val
1 5 10 15

Leu Glu Arg Leu Tyr Ala Ser Pro Met Asp Gly Thr Ser Ser Ala Leu
20 25 30

Ser Met Gly Pro Phe Val Pro Phe Ile Met Arg Cys Gly His Ser Pro
35 40 45

Val Tyr His Ser Arg Glu Met Ala Ala Arg Ala Leu Val Pro Phe Val
50 55 60

Met Ile Asp His Ile Pro Asn Thr Ile Arg Thr Leu Leu Ser Thr Leu
65 70 75 80

Pro Ser Cys Thr Asp Gln Cys Phe Arg Ala Lys Pro His Ser Trp Gly
85 90 95

His Phe Ser Arg Phe Phe His Leu Leu Gln Ala Tyr Ser Asp Ser Lys
100 105 110

Thr Arg Asn Glu Phe Arg Leu Pro Ala Arg Ala Asp
115 120

<210> 248

<211> 674

<212> PRT

<213> Homo sapiens

<400> 248

Met Thr Gly Arg Glu Phe Phe Ser Arg Phe Pro Glu Leu Tyr Pro Phe
1 5 10 15

Leu Leu Lys Gln Leu Glu Thr Val Ala Asn Thr Val Asp Ser Asp Met
20 25 30

Gly Glu Pro Asn Arg His Pro Ser Met Phe Leu Leu Leu Val Leu

		175
35	40	45
Glu Arg Leu Tyr Ala Ser Pro Met Asp Gly Thr Ser Ser Ala Leu Ser		
50	55	60
Met Gly Pro Phe Val Pro Phe Ile Met Arg Cys Gly His Ser Pro Val		
65	70	75
Tyr His Ser Arg Glu Met Ala Ala Arg Ala Leu Val Pro Phe Val Met		
85	90	95
Ile Asp His Ile Pro Asn Thr Ile Arg Thr Leu Leu Ser Thr Leu Pro		
100	105	110
Ser Cys Thr Asp Gln Cys Phe Arg Gln Asn His Ile His Gly Thr Leu		
115	120	125
Leu Gin Val Phe His Leu Leu Gln Ala Tyr Ser Asp Ser Lys His Gly		
130	135	140
Thr Asn Ser Asp Phe Gln His Glu Leu Thr Asp Ile Thr Val Cys Thr		
145	150	155
Lys Ala Lys Leu Trp Leu Ala Lys Arg Gln Asn Pro Cys Leu Val Thr		
165	170	175
Arg Ala Val Tyr Ile Asp Ile Leu Phe Leu Leu Thr Cys Cys Leu Asn		
180	185	190
Arg Ser Ala Lys Asp Asn Gln Pro Val Leu Glu Ser Leu Gly Phe Trp		
195	200	205
Glu Glu Val Arg Gly Ile Ile Ser Gly Ser Glu Leu Ile Thr Gly Phe		
210	215	220
Pro Trp Ala Phe Lys Val Pro Gly Leu Pro Gln Tyr Leu Gln Ser Leu		
225	230	235
Thr Arg Leu Ala Ile Ala Ala Val Trp Ala Ala Ala Lys Ser Gly		
245	250	255
Glu Arg Glu Thr Asn Val Pro Ile Ser Phe Ser Gln Leu Leu Glu Ser		
260	265	270
Ala Phe Pro Glu Val Arg Ser Leu Thr Leu Glu Ala Leu Leu Glu Lys		
275	280	285
Phe Leu Ala Ala Ala Ser Gly Leu Gly Glu Lys Gly Val Pro Pro Leu		
290	295	300
Leu Cys Asn Met Gly Glu Lys Phe Leu Leu Leu Ala Met Lys Glu Asn		
305	310	315
His Pro Glu Cys Phe Cys Lys Ile Leu Lys Ile Leu His Cys Met Asp		
325	330	335
Pro Gly Glu Trp Leu Pro Gln Thr Glu His Cys Val His Leu Thr Pro		
340	345	350

176

Lys Glu Phe Leu Ile Trp Thr Met Asp Ile Ala Ser Asn Glu Arg Ser
355 360 365

Glu Ile Gln Ser Val Ala Leu Arg Leu Ala Ser Lys Val Ile Ser His
370 375 380

His Met Gln Thr Cys Val Glu Asn Arg Glu Leu Ile Ala Ala Glu Leu
385 390 395 400

Lys Gln Trp Val Gln Leu Val Ile Leu Ser Cys Glu Asp His Leu Pro
405 410 415

Thr Glu Ser Arg Leu Ala Val Val Glu Val Leu Thr Ser Thr Thr Pro
420 425 430

Leu Phe Leu Thr Asn Pro His Pro Ile Leu Glu Leu Gln Asp Thr Leu
435 440 445

Ala Leu Trp Lys Cys Val Leu Thr Leu Leu Gln Ser Glu Glu Gln Ala
450 455 460

Val Arg Asp Ala Ala Thr Glu Thr Val Thr Thr Ala Met Ser Gln Glu
465 470 475 480

Asn Thr Cys Gln Ser Thr Glu Phe Ala Phe Cys Gln Val Asp Ala Ser
485 490 495

Ile Ala Leu Ala Leu Ala Leu Ala Val Leu Cys Asp Leu Leu Gln Gln
500 505 510

Trp Asp Gln Leu Ala Pro Gly Leu Pro Ile Leu Leu Gly Trp Leu Leu
515 520 525

Gly Glu Ser Asp Asp Leu Val Ala Cys Val Glu Ser Met His Gln Val
530 535 540

Glu Glu Asp Tyr Leu Phe Glu Lys Ala Glu Val Asn Phe Trp Ala Glu
545 550 555 560

Thr Leu Ile Phe Val Lys Tyr Leu Cys Lys His Leu Phe Cys Leu Leu
565 570 575

Ser Lys Ser Gly Trp Arg Pro Pro Ser Pro Glu Met Leu Cys His Leu
580 585 590

Gln Arg Met Val Ser Glu Gln Cys His Leu Leu Ser Gln Phe Phe Arg
595 600 605

Glu Leu Pro Pro Ala Ala Glu Phe Val Lys Thr Val Glu Phe Thr Arg
610 615 620

Leu Arg Ile Gln Glu Glu Arg Thr Leu Ala Cys Leu Arg Leu Leu Ala
625 630 635 640

Phe Leu Glu Gly Lys Glu Gly Glu Asp Thr Leu Val Leu Ser Val Trp
645 650 655

177

Asp Ser Tyr Ala Glu Ser Arg Gln Leu Thr Leu Pro Arg Thr Glu Ala
 660 665 670

Ala Cys

<210> 249
 <211> 10
 <212> PRT
 <213> Homo sapiens

<400> 249
 Ile Ile Ser Gly Ser Glu Leu Ile Thr Gly
 1 5 10

<210> 250
 <211> 230
 <212> PRT
 <213> Homo sapiens

<400> 250
 Val Asp Gly Ile Asp Lys Leu Asp Ile Glu Phe Leu Gln Gln Phe Leu
 1 5 10 15

Glu Thr His Ser Arg Gly Pro Arg Leu His Ser Pro Gly His Ala Ser
 20 25 30

Gln Glu Ala Thr Pro Gly Ala Asn Met Ser Ser Gly Thr Glu Leu Leu
 35 40 45

Trp Pro Gly Ala Ala Leu Leu Val Leu Leu Gly Val Ala Ala Ser Leu
 50 55 60

Cys Val Arg Cys Ser Arg Pro Gly Ala Lys Arg Ser Glu Lys Ile Tyr
 65 70 75 80

Gln Gln Arg Ser Leu Arg Glu Asp Gln Gln Ser Phe Thr Gly Ser Arg
 85 90 95

Thr Tyr Ser Leu Val Gly Gln Ala Trp Pro Gly Pro Leu Ala Asp Met
 100 105 110

Ala Pro Thr Arg Lys Asp Lys Leu Leu Gln Phe Tyr Pro Ser Leu Glu
 115 120 125

Asp Pro Ala Ser Ser Arg Tyr Gln Asn Phe Ser Lys Gly Ser Arg His
 130 135 140

Gly Ser Glu Glu Ala Tyr Ile Asp Pro Ile Ala Met Glu Tyr Tyr Asn
 145 150 155 160

Trp Gly Arg Phe Ser Lys Pro Pro Glu Asp Asp Asp Ala Asn Ser Tyr
 165 170 175

Glu Asn Val Leu Ile Cys Lys Gln Lys Thr Thr Glu Thr Gly Ala Gln
 180 185 190

178

Gln Glu Gly Ile Gly Gly Leu Cys Arg Gly Asp Leu Ser Leu Ser Leu
195 200 205

Ala Leu Lys Thr Gly Pro Thr Ser Gly Leu Cys Pro Ser Ala Ser Pro
210 . 215 220

Glu Glu Asp Glu Gly Ile
225 230

<210> 251
<211> 122
<212> PRT
<213> *Homo sapiens*

.<400> 251
Val Leu Trp Arg Glu Ala Ser Ala Leu Val Leu Ser Asn Arg Leu Ser
1 5 10 15

Ser Gly Leu Leu His Asp Leu Leu Leu Gln Pro Ala Ile His Ser Arg
20 25 30

Val Ser Leu Gln Arg Ser Arg Val Leu Ser Ala Met Lys His Val Leu
50 55 60

Asn	Leu	Tyr	Leu	Leu	Gly	Val	Val	Leu	Thr	Leu	Leu	Ser	Ile	Phe	Val
65					70					75					80

Arg Val Met Glu Ser Leu Glu Gly Leu Leu Glu Ser Pro Ser Pro Gly
85 90 95

Thr Ser Trp Thr Thr Arg Ser Gln Leu Ala Asn Thr Glu Pro Thr Lys
100 105 110

Gly Leu Pro Asp His Pro Ser Arg Ser Met
115 120

<210> 252
<211> 129
<212> PRT
<213> *Homo sapiens*

<400> 252
Tyr Thr Phe His Thr Gln Ile Phe Leu Asp Phe Pro Met Ile Phe Leu
1 5 10 15

Thr Val Leu Pro Leu Ala Phe Leu Phe Leu His Ser Gly Phe Tyr His
 20 25 30

Tyr Ile Ser Phe Ser Cys Leu Phe Ser Leu Ser Leu Ala Leu Phe Phe
35 40 45

Phe Leu Asp Val Ala Thr Phe Arg Arg Pro Gly Gln Leu Phe Cys Glu

179

50 55 60

Arg Ser Val Leu Phe Asp Met Phe His Phe Gly Phe Val Ser Leu Phe
65 70 75 80

Leu His Glu Trp Ile Gln Ala Lys His Phe Trp Ala Gly Leu Phe Ile
85 90 95

Val Leu Pro Ser Asp Val Phe Phe Ser Val His His Leu Glu Ala Pro
100 105 110

Asp Gly Ser Phe Pro Asn Ile Ala Lys Leu Ser Leu Ile Ile Leu Leu
115 120 125

Arg

<210> 253
<211> 99
<212> PRT
<213> *Homo sapiens*

<400> 253
Gly Thr Arg Phe Pro Thr Gly Glu Thr Pro Ser Leu Gly Phe Thr Val
1 5 10 15

Thr Leu Val Leu Leu Asn Ser Leu Ala Phe Leu Leu Met Ala Val Ile
20 25 30

Tyr Thr Lys Leu Tyr Cys Asn Leu Glu Lys Glu Asp Leu Ser Glu Asn
 35 40 45

Ser Gln Ser Ser Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn
50 55 60

Cys Ile Phe Phe Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile
65 70 75 80

Thr Ala Ile Ser Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile
85 90 95

Phe Phe Pro

<210> 254
<211> 51
<212> PRT
<213> *Homo sapiens*

<400> 254

Met	Ile	Lys	His	Val	Ala	Trp	Leu	Ile	Phe	Thr	Asn	Cys	Ile	Phe	Phe
1				5					10					15	

Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile Thr Ala Ile Ser
20 25 30

180

Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile Phe Phe Pro Cys
 35 40 45

Leu Leu Ala
 50

<210> 255
 <211> 259
 <212> PRT
 <213> Homo sapiens

<400> 255
 Gly Thr Arg Phe Pro Thr Gly Glu Thr Pro Ser Leu Gly Phe Thr Val
 1 5 10 15

Thr Leu Val Leu Leu Asn Ser Leu Ala Phe Leu Leu Met Ala Val Ile
 20 25 30

Tyr Thr Lys Leu Tyr Cys Asn Leu Glu Lys Glu Asp Leu Ser Glu Asn
 35 40 45

Ser Gln Ser Ser Met Ile Lys His Val Ala Trp Leu Ile Phe Thr Asn
 50 55 60

Cys Ile Phe Phe Cys Pro Val Ala Phe Phe Ser Phe Ala Pro Leu Ile
 65 70 75 80

Thr Ala Ile Ser Ile Ser Pro Glu Ile Met Lys Ser Val Thr Leu Ile
 85 90 95

Phe Phe Pro Leu Pro Ala Cys Leu Asn Pro Val Leu Tyr Val Phe Phe
 100 105 110

Asn Pro Lys Phe Lys Glu Asp Trp Lys Leu Leu Lys Arg Arg Val Thr
 115 120 125

Lys Lys Ser Gly Ser Val Ser Val Ser Ile Ser Ser Gln Gly Gly Cys
 130 135 140

Leu Glu Gln Asp Phe Tyr Tyr Asp Cys Gly Met Tyr Ser His Leu Gln
 145 150 155 160

Gly Asn Leu Thr Val Cys Asp Cys Cys Glu Ser Phe Leu Leu Thr Lys
 165 170 175

Pro Val Ser Cys Lys His Leu Ile Lys Ser His Ser Cys Pro Ala Leu
 180 185 190

Ala Val Ala Ser Cys Gln Arg Pro Glu Gly Tyr Trp Ser Asp Cys Gly
 195 200 205

Thr Gln Ser Ala His Ser Asp Tyr Ala Asp Glu Glu Asp Ser Phe Val
 210 215 220

Ser Asp Ser Ser Asp Gln Val Gln Ala Cys Gly Arg Ala Cys Phe Tyr
 225 230 235 240

181

Gln Ser Arg Gly Phe Pro Leu Val Arg Tyr Ala Tyr Asn Leu Pro Arg
245 250 255

Val Lys Asp

<210> 256
<211> 22
<212> PRT
<213> *Homo sapiens*

<400> 256
Cys Asp Cys Cys Glu Ser Phe Leu Leu Thr Lys Pro Val Ser Cys Lys
1 5 10 15

His Leu Ile Lys Ser His
20

<210> 257
<211> 81
<212> PRT
<213> *Homo sapiens*

<220>
<221> SITE
<222> (20)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 257
Ala Leu Glu Asn Ser Gly Ser Pro Gly Leu Gln Asp Ser Ala Arg Ala
1 5 10 15

His Phe Asn Xaa Ser Leu Arg Ser Phe Ser Phe Leu Arg Asn Gln Met
20 25 30

Tyr Ile Phe Glu Leu Ser Leu Tyr Leu Glu Gly Thr Ser Phe Val Val
35 40 45

Val Leu Leu Phe Leu Leu Ile Ser Val Ser Leu Asp Ser Pro Pro Thr
 50 55 60 .

Thr Lys Gly Trp Asp Ser Val Leu His Ile Trp Val Pro Leu Ile Val
 65 70 75 80

G1n

<210> 258
<211> 77
<212> PRT
<213> *Homo sapiens*

<400> 258
Gly His Glu Ser Ile Cys Gly Ser Cys Arg Ser Trp Ile Tyr Phe Ser
1 5 10 15

182

Ile	Arg	Cys	Arg	Arg	Arg	Met	Arg	Pro	Trp	Trp	Ser	Leu	Leu	Leu	Glu
20														30	

Ala	Cys	Ala	Thr	Cys	Ala	Gln	Thr	Gly	Pro	Thr	Arg	Ser	Thr	Ser	Cys
35														45	

Thr	Gln	Glu	Val	Ser	His	Ser	Ser	Ser	Thr	Ala	Tyr	Pro	Ala	Pro	Met
50														60	

Arg	Arg	Arg	Cys	Cys	Leu	Pro	Ser	Pro	Arg	Ser	Cys	Thr			
65													75		

<210> 259

<211> 119

<212> PRT

<213> Homo sapiens

<400> 259

Lys	Arg	Ala	Gly	Val	Glu	Val	Gly	Gly	Leu	Val	Met	Ala	Leu	Ala	Gly
1														15	

Ser	Val	Phe	Val	Leu	Gly	Gly	Val	Leu	Val	Leu	Cys	Val	Glu	Arg	Asn
20														30	

Gly	Glu	Gly	Glu	Met	Gly	Trp	Pro	Gln	His	Leu	Pro	Lys	Ser	Gln	Pro
35														45	

Leu	Ser	Pro	Pro	Val	Ala	Val	Arg	Arg	Cys	Ser	Phe	Glu	Arg	Ser	Trp
50														60	

Ile	Asp	Leu	Leu	Val	Glu	Thr	Ser	Ser	Ser	Met	Val	Thr	Cys	Arg	Gln
65														80	

Gln	Val	Gly	Thr	Pro	Asn	Gly	Met	Glu	Gly	Arg	Gly	Gly	Pro	Lys	
85														95	

Thr	Thr	Phe	Pro	Ile	Arg	Leu	Gln	Leu	Ser	Gly	Ala	Cys	Ala	Val	Arg
100														110	

Pro	Glu	Ile	Gln	Trp	Glu	Val									
							115								

<210> 260

<211> 275

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (47)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (94)

183

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (192)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 260

Gln Asp Trp Lys Ala Glu Arg Ser Gln Asp Pro Phe Glu Lys Cys Met
1 5 10 15Gln Asp Pro Asp Tyr Glu Gln Leu Leu Lys Val Thr Ile Leu Glu Ala
20 25 30Asp Asn Arg Ile Gly Gly Arg Ile Phe Thr Tyr Arg Asp Gln Xaa Thr
35 40 45Gly Trp Ile Gly Glu Leu Gly Ala Met Arg Met Pro Ser Ser His Arg
50 55 60Ile Leu His Lys Leu Cys Gln Gly Leu Gly Leu Asn Leu Thr Lys Phe
65 70 75 80Thr Gln Tyr Asp Lys Asn Thr Trp Thr Glu Val His Glu Xaa Lys Leu
85 90 95Arg Asn Tyr Val Val Glu Lys Val Pro Glu Lys Leu Gly Tyr Ala Leu
100 105 110Arg Pro Gln Glu Lys Gly His Ser Pro Glu Asp Ile Tyr Gln Met Ala
115 120 125Leu Asn Gln Ala Leu Lys Asp Leu Lys Ala Leu Gly Cys Arg Lys Ala
130 135 140Met Lys Lys Phe Glu Arg His Thr Leu Leu Glu Tyr Leu Leu Gly Glu
145 150 155 160Gly Asn Leu Ser Arg Pro Ala Val Gln Leu Leu Gly Asp Val Met Ser
165 170 175Glu Asp Gly Phe Phe Tyr Leu Ser Phe Ala Glu Ala Leu Arg Ala Xaa
180 185 190Ser Cys Leu Ser Asp Arg Leu Gln Tyr Ser Arg Ile Val Gly Gly Trp
195 200 205Asp Leu Leu Pro Arg Ala Leu Leu Ser Ser Leu Ser Gly Leu Val Leu
210 215 220Leu Asn Ala Pro Val Val Ala Met Thr Gln Gly Pro His Asp Val His
225 230 235 240Val Gln Ile Glu Thr Ser Pro Pro Ala Arg Asn Leu Lys Val Leu Lys
245 250 255Ala Asp Val Val Leu Leu Thr Ala Ser Gly Pro Ala Val Lys Arg Ile
260 265 270

Thr Phe Ser
275

<210> 261
<211> 212
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (123)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 261
Leu Pro Arg His Met Gln Glu Ala Leu Arg Arg Leu His Tyr Val Pro
1 5 10 15

Ala Thr Lys Val Phe Leu Ser Phe Arg Arg Pro Phe Trp Arg Glu Glu
20 25 30

His Ile Glu Gly Gly His Ser Asn Thr Asp Arg Pro Ser Arg Met Ile
35 40 45

Phe Tyr Pro Pro Pro Arg Glu Gly Ala Leu Leu Leu Ala Ser Tyr Thr
50 55 60

Trp Ser Asp Ala Ala Ala Phe Ala Gly Leu Ser Arg Glu Glu Ala
65 70 75 80

Leu Arg Leu Ala Leu Asp Asp Val Ala Ala Leu His Gly Pro Val Val
85 90 95

Arg Gln Leu Trp Asp Gly Thr Gly Val Val Lys Arg Trp Ala Glu Asp
100 105 110

Gln His Ser Gln Gly Gly Phe Val Val Gln Xaa Pro Ala Leu Trp Gln
115 120 125

Thr Glu Lys Asp Asp Trp Thr Val Pro Tyr Gly Arg Ile Tyr Phe Ala
130 135 140

Gly Glu His Thr Ala Tyr Pro His Gly Trp Val Glu Thr Ala Val Lys
145 150 155 160

Ser Ala Leu Arg Ala Ala Ile Lys Ile Asn Ser Arg Lys Gly Pro Ala
165 170 175

Ser Asp Thr Ala Ser Pro Glu Gly His Ala Ser Asp Met Glu Gly Gln
180 185 190

Gly His Val His Gly Val Ala Ser Ser Pro Ser His Asp Leu Ala Lys
195 200 205

Glu Glu Gly Ser
210

<210> 262
<211> 319
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (68)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (115)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (213)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 262
Met Ala Pro Leu Ala Leu His Leu Leu Val Leu Val Pro Ile Leu Leu
1 5 10 15

Ser Leu Val Ala Ser Gln Asp Trp Lys Ala Glu Arg Ser Gln Asp Pro
20 25 30

Phe Glu Lys Cys Met Gln Asp Pro Asp Tyr Glu Gln Leu Leu Lys Val
35 40 45

Thr Ile Leu Glu Ala Asp Asn Arg Ile Gly Gly Arg Ile Phe Thr Tyr
50 55 60

Arg Asp Gln Xaa Thr Gly Trp Ile Gly Glu Leu Gly Ala Met Arg Met
65 70 75 80

Pro Ser Ser His Arg Ile Leu His Lys Leu Cys Gln Gly Leu Gly Leu
85 90 95

Asn Leu Thr Lys Phe Thr Gln Tyr Asp Lys Asn Thr Trp Thr Glu Val
100 105 110

His Glu Xaa Lys Leu Arg Asn Tyr Val Val Glu Lys Val Pro Glu Lys
115 120 125

Leu Gly Tyr Ala Leu Arg Pro Gln Glu Lys Gly His Ser Pro Glu Asp
130 135 140

Ile Tyr Gln Met Ala Leu Asn Gln Ala Leu Lys Asp Leu Lys Ala Leu
145 150 155 160

Gly Cys Arg Lys Ala Met Lys Lys Phe Glu Arg His Thr Leu Leu Glu
165 170 175

Tyr Leu Leu Gly Glu Gly Asn Leu Ser Arg Pro Ala Val Gln Leu Leu
180 185 190

186

Gly Asp Val Met Ser Glu Asp Gly Phe Phe Tyr Leu Ser Phe Ala Glu
195 200 205

Ala Leu Arg Ala Xaa Ser Cys Leu Ser Asp Arg Leu Gln Tyr Ser Arg
210 215 220

Ile Val Gly Gly Trp Asp Leu Leu Pro Arg Ala Leu Leu Ser Ser Leu
225 230 235 240

Ser Gly Leu Val Leu Leu Asn Ala Pro Val Val Ala Met Thr Gln Gly
245 250 255

Pro His Asp Val His Val Gln Ile Glu Thr Ser Pro Pro Ala Arg Asn
260 265 270

Leu Lys Val Leu Lys Ala Asp Val Val Leu Leu Thr Ala Ser Gly Pro
275 280 285

Ala Val Lys Arg Ile Thr Phe Ser Pro Arg Cys Pro Ala Thr Cys Arg
290 295 300

Arg Arg Cys Gly Gly Cys Thr Thr Cys Arg Pro Pro Arg Cys Ser
305 310 315

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/15849

A. CLASSIFICATION OF SUBJECT MATTER IPC(6) :C12N 15/11, 15/00, 15/63; C07H 21/02, 21/04 US CL : 536/23.1, 23.4; 435/320.1, 69.1 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 536/23.1, 23.4; 435/320.1, 69.1		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Sequence searched SEQ ID NO:11, SEQ ID NO: 103, APS, STN, CAPLUS, terms nucleic acid, express, vector, pancreas islet cell tumors.		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A, P	US 5,849,498 A (BANDMAN et al) 15 December 1998, sequence listing.	1-10, 14-15, and 21
A	US 5,670,367 A (DORNER et al) 23 September 1997, especially sequence listing.	1-10, 14-15, and 21
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <ul style="list-style-type: none"> *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed <p>*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>*A* document member of the same patent family</p>		
Date of the actual completion of the international search 30 SEPTEMBER 1999	Date of mailing of the international search report 21 OCT 1999	
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	<p>Authorized officer <i>D. Lawrence Lee</i> LI LEE</p> <p>Telephone No. (703) 308-0196</p>	

INTERNATIONAL SEARCH REPORT

International application No. PCT/US99/15849

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-10, 14-15, and 21

Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/15849

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s)1-10, 14-15, and 21, drawn to isolated nucleic acid and expression system.

Group II, claim(s) 11-12 and 16, drawn to isolated polypeptide.

Group III, claim(s) 13, drawn to antibody.

Group IV, claim(s)17, drawn to method for preventing a medical condition.

Group V, claim(s) 18-19, drawn to method of diagnosing a disease.

Group VI, claim(s) 20, 22, drawn to method for identifying a binding partner.

Group VII, claim 23, drawn to product produced by method of claim 20.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows:

The inventions listed as Groups I-VII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The special technical feature of Group I is considered to be isolated nucleic acid and expression system.

The special technical feature of Group II is considered to be isolated polypeptide.

The special technical feature of Group III is considered to be antibody.

The special technical feature of Group IV is considered to be a method for preventing a medical condition..

The special technical feature of Group V is considered to be method of diagnosing a disease.

The special technical feature of Group VI is considered to be a method for identifying a binding partner..

The special technical feature of Group VII is considered to be product produced by method of claim 20.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows:

There are 71 genes, from gene Nos 1-71.

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons: The 71 genes have different nucleic acid sequences and they are from different cell. Therefore they lack same special technical features.